Universidade Federal Fluminense
Centro Tecnológico
Escola de Engenharia
Mestrado em Engenharia Elétrica e de Telecomunicações

Diego Santana Marques Bahiano

Criação de modelos estocásticos para a síntese de vogais considerando os pulsos glotais de Rosenberg e de Liljencrants-Fant com parâmetros unificados

Niterói-RJ
2020

DIEGO SANTANA MARQUES BAHIANO

Criação de modelos estocásticos para a síntese de vogais considerando os pulsos glotais de Rosenberg e de Liljencrants-Fant com parâmetros unificados

Dissertação apresentada ao Curso de Mestrado em Engenharia Elétrica e de Telecomunicações da Universidade Federal Fluminense, como requisito parcial para obtenção do Grau de Mestre. Área de Concentração: Sistemas de Telecomunicações.

Orientador: Prof. D.Sc. EDSON LUIZ CATALDO FERREIRA

Niterói-RJ

2020

Ficha catalográfica automática - SDC/BEE
Gerada com informações fornecidas pelo autor

B151c Bahiano, Diego Santana Marques
Criação de modelos estocásticos para a síntese de vogais considerando os pulsos glotais de Rosenberg e de Liljencrants-Fant com parâmetros unificados / Diego Santana Marques Bahiano ; Edson Luiz Cataldo Ferreira, orientador. Niterói, 2020.
94 f. : il.

Dissertação (mestrado)-Universidade Federal Fluminense, Niterói, 2020.

DOI: http://dx.doi.org/10.22409/PPGEET.2020.m.85260479572

1. Processamento Digital de Sinais. 2. Processos Estocásticos. 3. Produção de voz. 4. Produção intelectual. I. Ferreira, Edson Luiz Cataldo, orientador. II. Universidade Federal Fluminense. Escola de Engenharia. III. Título.

CDD -

Bibliotecário responsável: Sandra Lopes Coelho - CRB7/3389

DIEGO SANTANA MARQUES BAHIANO

Criação de modelos estocásticos para a síntese de vogais considerando os pulsos glotais de Rosenberg e de Liljencrants-Fant com parâmetros unificados

Dissertação apresentada ao Curso de Mestrado em Engenharia Elétrica e de Telecomunicações da Universidade Federal Fluminense, como requisito parcial para obtenção do Grau de Mestre. Área de Concentração: Sistemas de Telecomunicações.

Aprovada em 14 de maio de 2020.

BANCA EXAMINADORA

Prof. D.Sc. EDSON LUIZ CATALDO FERREIRA - Orientador
UFF

Profa. D.Sc. LENI JOAQUIM DE MATOS
UFF

Prof. D.Sc. THIAGO GAMBOA RITTO
UFRJ

Niterói-RJ
2020

Dedicado a Deus e aos meus pais.

Agradecimentos

Agradeço a Deus pela vida e pelo destino lindo ao qual tem me direcionado.

Agradeço aos meus pais e dois irmãos por todo amor que me transmitem diariamente, mesmo que por vezes não presencialmente, sem o qual nada seria realidade.

Agradeço aos meus amigos, os quais não listarei para não correr o risco de esquecer algum nome. Tenho a certeza de que os levarei sempre comigo e sei que eles continuarão tornando a minha vida ainda mais extraordinária, inigualável e menos árdua.

Agradeço em especial ao professor Edson Cataldo pela orientação, prestatividade a todo tempo durante a construção deste trabalho, conhecimento, confiança, apoio e, sobretudo, pela paciência em me ajudar a encontrar soluções eficientes para os problemas encontrados.

Agradeço a todos os meus docentes da UFF e da UFBA, em especial aos professores Natália e Murilo, os quais foram de fundamental importância para construção do conhecimento necessário para desenvolvimento desta dissertação.

Agradeço aos funcionários da secretaria do programa de pós-graduação e, em especial, à Vânia pela solicitude e gentileza.

Agradeço à Universidade Federal Fluminense e ao PPGEET por me conceder a incrível oportunidade de me tornar Mestre.

Lista de Figuras

Lista de Tabelas

Sumário

Resumo

Na produção da voz, a variação dos comprimentos dos ciclos glotais (ciclo completo das cordas vocais desde o início de sua abertura até seu fechamento completo) em relação a um valor médio, causada pelo movimento (quase) periódico das cordas vocais, gera o fenômeno aleatório conhecido por *jitter*. Seu estudo é importante devido a aplicações como a identificação de patologias relacionadas à voz, melhora da naturalidade da voz sintetizada, simulação de vozes com características de patologias e, também, favorecimento na calibragem de algoritmos de processamento de sinais para identificação de ciclos glotais. Esse trabalho objetiva construir modelos estocásticos para o sinal glotal, baseados em dois modelos determinísticos conhecidos da literatura, o modelo de Rosenberg e o modelo de Liljencrantes-Fant com parâmetros unificados, considerando o intervalo de tempo glotal como um processo estocástico. Consequentemente, sinais glotais com *jitter* serão obtidos e, usando a teoria fonte-filtro de Fant, serão sintetizados sinais de vozes normais e, também, com características de patologias. Todos os algoritmos foram desenvolvidos em Python.

Palavras-chave: *jitter*, processo estocástico, sinal glotal, síntese de voz.

Abstract

In voice production, the variation of the glottal cycles lengths (complete cycle of the vocal folds since their opening up to the complete closing) in relation to a mean value, caused by the (quasi) periodic movement of the vocal folds, generates the random phenomenon known as *jitter*. Its study is important due to applications as the identification of voice-related pathologies, the improvement of the synthesized voice naturalness, simulation of voices with pathological characteristics and the generation of voice signals to calibrate signal processing algorithms and to help in detecting glottal cycles, as well. This work aims to build stochastic models for the glottal signal, based two deterministic models known from the literature, the Rosenberg model and the Liljencrants-Fant model with unified parameters, considering the glottal time interval as a stochastic process. Consequently, glottal signal with *jitter* are obtained and, using the Fant source-filter theory, normal voice signals and with pathological characteristics are synthesized. All the algorithms were developed in Python.

Keywords: *jitter*, stochastic process, glottal signal, speech synthesis.

Capítulo 1

Introdução

O ser humano é capaz de expressar seus sentimentos, pensamentos e vontades através da voz, sendo esta um meio de comunicação de grande importância para a vida em sociedade. O aparelho fonador é constituído pelo aparelho respiratório, pela fonte de vibração (localizada na laringe) e pelo trato vocal (composto por faringe, boca e nariz). O fluxo de ar proveniente dos pulmões, após passar pelas cordas vocais, torna-se um sinal de pressão acústica (quase) periódico que irá se propagar pelo trato vocal, sendo filtrado e amplificado, até ser irradiado pela boca gerando a voz [1].

Com o avanço tecnológico e a necessidade do estudo das teorias e fenômenos associados à fala, torna-se de suma importância a análise das muitas variáveis relacionadas com a produção da voz, tais como a frequência fundamental do som, as medidas de perturbação (*jitter* e *shimmer*) e os ruídos. Dessa forma, a construção de modelos estocásticos que expressem da melhor forma o sistema acústico fonador é um grande desafio para os estudiosos da área.

A especificidade dos órgãos envolvidos na produção da voz torna o sistema fonador único e injetivo para cada indivíduo e, sendo assim, o estudo envolvendo os fatores do processo, desde o fluxo de ar gerado pelos pulmões até a configuração de sons audíveis, é bastante relevante para a sua modelagem. É possível identificar neste contexto algumas variáveis que estão, inclusive, associadas a patologias das cordas vocais, o que é o caso do fenômeno *jitter*, que diz respeito às pequenas flutuações dos períodos do pulso glotal gerado.

Medidas de *jitter* local, que estão relacionadas às variações do período fundamental do sinal glotal, superiores a 1% é um indício da existência de patologias nas cordas vocais [2]. Dessa forma, a análise dessa característica da voz permite a construção de modelos que conseguem dar maior naturalidade ao som sintetizado, isto aliado à escolha de formantes, bandas de passagem e demais variáveis que são definidas no sistema, conforme o tipo de som que se deseja reproduzir. Também possibilita simular vozes roucas e/ou com características de patologias e, ainda, podem ajudar a calibrar sistemas de processamento de sinais de voz que necessitem identificar ciclos glotais.

Há alguns estudos voltados à construção de modelos para a produção de voz e, para atingir este objetivo, é indispensável o conhecimento das propriedades do *jitter*, o qual confere maior semelhança ao som produzido pelos humanos e será abordado no capítulo IV, por ser um fenômeno intrínseco ao aparelho fonador. Além deste fato, outras motivações são levadas em conta para o atual investimento nesta área de pesquisa, a exemplo da constatação do aumento do *jitter* em vozes disfônicas; a necessidade de implementação de modelo para simular vozes mais naturais ou imitar vozes roucas; e a melhoria de algoritmos que identificam ciclos glotais e seus comprimentos.

Schoentgen apresentou em [3] modelos de *jitter* que simulavam os distúrbios nas frequências instantâneas de vibração das paredes glotais causando perturbações nos comprimentos dos ciclos glotais. A modelagem combinou um modelo estocástico das perturbações com o chamado modelo de fita de vibração das cordas vocais, modelo este que levou em consideração microtremores e ruídos livres de correlação. Foi observado, através da experimentação, que patologias da laringe alteram indiretamente o *jitter* através de uma variação na resposta das vibrações glotais aos distúrbios externos.

Em [4], é apresentado um algoritmo de modificação de voz para transformar uma voz normal numa voz rouca. O algoritmo é baseado na modificação da tremulação da voz, o que está relacionado à rouquidão, fenômeno associado à mudança no tom ou na qualidade da voz, em geral para um tom mais áspero. Um processo de predição linear é usado para aumentar o *jitter* e o brilho nas vozes, das quais os dados para o *jitter* e o *shimmer* são obtidos através de um modelo estocástico baseado em vozes naturalmente roucas. A avaliação formal do algoritmo realizado com múltiplas vozes e ouvintes mostrou que as vozes produzidas são percebidas o mais próximo possível da rouquidão natural, sugerindo a aplicabilidade do algoritmo para várias modificações e sínteses de voz.

Já em [2], é possível verificar a implementação de um diferente modelo estocástico para a criação do *jitter* no sistema, desenvolvido a partir do modelo mecânico massa-mola, o qual simula o movimento das cordas vocais na laringe. Foi possível no estudo proposto considerar a rigidez da mola como um processo estocástico com densidade espectral de potência definida e dependente de duas variáveis, as quais foram sendo modificadas a fim de simular a geração de vozes com e sem indícios de patologias.

Com base no modelo mecânico anterior, foi implementado outro processo estocástico mais complexo envolvendo três parâmetros de controle em [5], a partir do qual foram obtidos sons com diferentes níveis de *jitter*. Foi possível, nesta etapa, modificar inclusive a frequência fundamental da voz gerada, isto através da variação de determinados parâmetros, o que também proporcionou estabelecer o controle do colorido espectral do sistema. Esta dissertação foi desenvolvida a partir dessas ideias.

Fora os autores já citados, ainda há muitas outras pesquisas acerca do fenômeno *jitter*, contribuindo, assim, para o desenvolvimento científico. Segundo [6], a maioria das patologias relacionadas à voz é devida a massas irregulares localizadas nas pregas vocais, interferindo em sua normalidade e vibração regular, o que diminui a qualidade da voz, sendo este geralmente o primeiro sintoma desse tipo de distúrbio. Vale ressaltar que, no passado, a única maneira de medir a qualidade da voz era aplicando medições perceptivas, as quais acusavam a existência ou ausência de várias características próprias da voz [7]. Assim, hoje, o *jitter* tem surgido como uma boa aposta para diagnósticos médicos com alto grau de sensibilidade/especificidade.

Há uma crescente demanda no desenvolvimento de técnicas que possam avaliar a qualidade da voz de maneira objetiva, fornecendo uma medição robusta e confiável de importantes parâmetros acústicos [8]. Dessa forma, com o recente desenvolvimento de tecnologias e a busca por equipamentos de aferição da qualidade da voz cada vez mais sofisticados, já há alguns *softwares* para análise do sinal de fala, a fim de estimar vários parâmetros que indicam perturbações de amplitude e frequência, o nível de fuga de ar, o seu grau de turbulência, dentre outros, ferramentas estas para análise em tempo real que fornecem um *feedback* importante sobre o desempenho da voz do indivíduo.

A análise dos sistemas humanos são de alta complexidade e, a partir desta constatação, percebe-se que são inúmeros os desafios para a construção de modelos mais fidedignos e que representem as múltiplas nuances do sistema fonador humano. O objetivo deste trabalho é a implementação, a partir dos modelos determinísticos de pulso glotal de Rosenberg e de Liljencrants-Fant, com parâmetros unificados, dois diferentes processos estocásticos para o período fundamental instantâneo de cada ciclo, visando conceber vogais sustentadas sintéticas semelhantes às naturais produzidas por um indivíduo, com e sem características patológicas.

Para a implementação dos referidos modelos, é importante a definição da densidade espectral de potência, característica de grande relevância em estudos de sistemas estocásticos, uma vez que permite conhecer como a potência do sinal se distribui na frequência. A partir de então, foram fixados determinados parâmetros, identificadas as variáveis intimamente relacionadas à implementação do *jitter*, geradas as vogais com e sem este fenômeno e, com a finalidade de comparar os sons produzidos, foram feitas análises da sensibilidade das variáveis associadas aos processos estabelecidos.

A estrutura desta dissertação encontra-se dividida em sete capítulos e um conjunto de apêndices, que indicam os algoritmos na linguagem de programação *Python* que foram utilizados nas simulações. Os capítulos são apresentados a seguir:

1. O capítulo II mostra as características morfológicas dos sistemas envolvidos na produção da voz humana, indicando as atribuições dos órgãos envolvidos e a devida importância destes;

2. O capítulo III introduz a modelagem fonte-filtro para síntese de voz, bem como mostra os modelos determinísticos de pulsos glotais de Rosenberg e Liljencrants-Fant (LF);

3. O capítulo IV define o fenômeno *jitter* e traz a representação matemática das diversas medidas utilizadas para representá-lo;

4. O capítulo V expõe os modelos estocásticos utilizados para a construção dos pulsos glotais, a fim de que o efeito *jitter* seja inserido no sistema;

5. O capítulo VI desenvolve simulações, em Python, e faz a análise da sensibilidade das variáveis envolvidas nos sistemas; e

6. O capítulo VII aborda os resultados e sugere trabalhos futuros que podem se utilizar dos modelos propostos.

Capítulo 2

A voz humana

De uma forma geral, a voz humana consiste no som gerado pelos seres humanos usando suas cordas vocais para se comunicar com o mundo exterior, sendo um canal de elevada importância para a vida em sociedade, pela qual se pode expressar opiniões, informar fatos, acordar decisões e até mesmo sorrir. É um dos primitivos instrumentos musicais, tendo características específicas para cada indivíduo, o que permite a criação de ferramentas que reconheçam os locutores através do som produzido.

No contexto fisiológico, a voz pode ser vista como a mútua cooperação entre sistemas distintos, objetivando o alvo comum: um som inteligível [9]. Assim sendo, vale aqui ressaltar os três sistemas mais relevantes para a sua construção: o articulatório, o fonatório e o respiratório, os quais compõem juntos o aparelho fonador, desenhado na Fig. 2.1 . Quando se produz o som da fala, há a produção de uma corrente de ar que sai dos pulmões e vai até os lábios, passando por diversos órgãos e estruturas, os quais adicionam características biométricas ao mesmo.

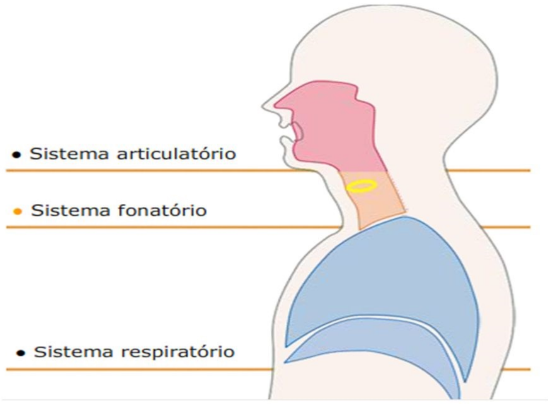


Figura 2.1: Aparelho fonador dividido em seus três sistemas [10].

Esses três sistemas citados apresentam particularidades que lhes conferem papéis indispensáveis para a fala. A fim de que se possa, ao final da leitura, aprender de forma significativa como uma passagem de ar por um duto humano é capaz de construir um sinal indispensável para o nascimento e desenvolvimento da sociedade, os três sistemas são descritos a seguir.

2.1 Articulatório

O sistema articulatório reúne as articulações do corpo, que podem ser classificadas em três grandes grupos: fibrosas, cartilaginosas e sinoviais. Representa a união entre duas ou mais estruturas rígidas (ossos, cartilagem e dente) que pode ou não apresentar movimento. No que tange o aparelho fonador, as estruturas que se localizam na parte superior da glote [11] que lhe pertencem são citadas, a seguir, com as devidas funções que desempenham.

2.1.1 Língua

É um órgão sensorial e muscular que está localizado na cavidade oral e faríngea, caracterizado por ser recoberto por uma membrana com papilas gustativas em sua superfície e os corpúsculos de Krause, responsáveis por perceber as sensações táteis. A língua, no que se refere às duas funções que desempenha, a muscular é a responsável pela articulação das palavras durante a construção da linguagem falada. Para geração de vogais, o som passa direto pela boca sem necessidade do movimento gesticuloso da língua.

2.1.2 Faringe

É um órgão tubular que possui comunicação com o esôfago, os ouvidos, as fossas nasais e a laringe. Possui uma função importante regulada pela epiglote, a qual se movimenta a fim de permitir a passagem de ar para laringe ou de alimentos para o esôfago. No que tange à voz, ele permite que haja passagem do ar oriundo do pulmão e pode ser modelado como um tubo de seção circular com área variável.

2.1.3 Nariz

O nariz é uma estrutura saliente, vultosa e com formato variável, situado na linha média da face e com função de recepção e eliminação do ar pelo sistema respiratório através de narinas, sendo formado por uma estrutura de cartilagem hialina. É uma das caixas de ressonância, além da boca, que amplifica o som, sendo a região onde se produzem os sons nasais. Um buraco no céu da boca, na maioria das vezes gerado por fenda palatina, ou uma falha na musculatura pode fazer com que a voz saia sempre pelo nariz, deixando a voz nasalada.

2.1.4 Lábios

Os lábios são semimucosas, bordas da mucosa que reveste a boca humana. Suas características estão intimamente relacionadas com a etnia e os genes herdados. Usá-los para sugar na amamentação é uma das primeiras habilidades dominadas pelo ser humano na infância. Para além disso, representa um dos muitos pontos de articulação que ajudam a bloquear o ar que vem dos pulmões e formar fonemas.

2.1.5 Dentes

O dente é uma estrutura dura, saliente e esbranquiçada na boca, formada por polpa, dentina e esmalte, que é originada no maxilar e na arcada dentária, no caso do ser humano, sendo usada para cortar, prender e triturar os alimentos, preparando-os para serem deglutidos. No que tange à voz, como função secundária, a utilização dos dentes na articulação é de tal importância, que qualquer alteração ou deformidade poderá afetar a dicção e pronúncia de alguns fonemas, influenciando ou até prejudicando a comunicação.

2.2 Fonatório

O referido sistema é constituído basicamente pela laringe, onde se encontram as duas pregas vocais responsáveis pelos movimentos vibratórios sobre o ar que é expelido dos pulmões, definindo a frequência fundamental da voz gerada pelo indivíduo [9]. Cartilagens, músculos e membranas são as principais estruturas que o compõem, estando o mesmo localizado na região da garganta entre a traqueia e a faringe, da qual é separado pela epiglote.

A laringe se divide em três compartimentos: subglote, glote e supraglote. A mucosa da laringe forma dois pares de pregas: o primeiro par superior constitui as falsas cordas vocais ou pregas vestibulares. O segundo par inferior forma as cordas vocais verdadeiras, as quais estão representadas na Fig. 2.2. Quando o ar passa pela laringe, os músculos podem se contrair, modificando a abertura das cordas vocais e produzindo sons [12].

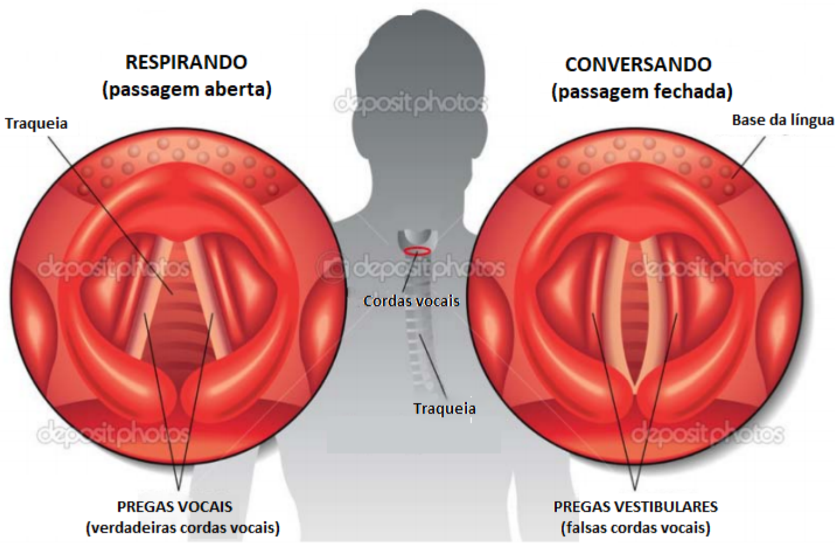


Figura 2.2: À esquerda: as cordas vocais durante a respiração, relaxadas e abertas. À direita: as cordas vocais formando uma pequena fenda durante a fala [13].

É devido à anatomia das cordas vocais que as vozes masculinas são mais graves e as femininas são mais agudas. No homem, as pregas vocais são mais grossas e mais elásticas; na mulher, as pregas são mais finas e tensas e, como consequência, vibram com maior frequência. As crianças possuem vozes agudas porque as suas cordas vocais são ainda mais curtas. Vale salientar que quando chega a puberdade, a voz fica mais grave nos meninos porque o hormônio masculino, além de promover inúmeras transformações no corpo, aumenta a massa das pregas vocais. Com o passar das décadas, com o processo natural de envelhecimento, ocorrem atrofia e arqueamento das pregas vocais e a musculatura da laringe fica mais flácida, o que torna a voz trêmula, rouca, soprosa, com pouca projeção [14].

2.3 Respiratório

O sistema respiratório é responsável pela hematose pulmonar, mecanismo químico-molecular indispensável para a oxigenação celular. No tocante à produção da voz, os órgãos principais [11] relacionados serão listados a seguir e indicados na Fig. 2.3.

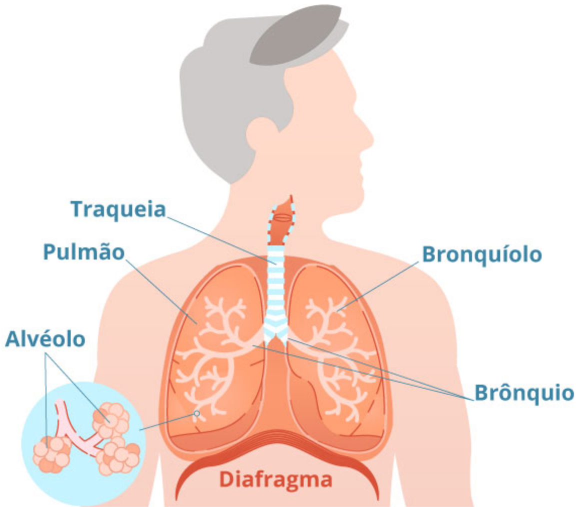


Figura 2.3: Principais órgãos do sistema respiratório associados à produção da voz [15].

2.3.1 Traqueia

A função principal desse órgão é dar um tratamento ao ar que vai até os pulmões, aquecendo-o, umidificando-o, filtrando-o e retendo as partículas mais sólidas, poeiras e microorganismos através dos cílios e mucos que lhe recobrem. Está localizada logo após a laringe e tem o formato de um tubo vertical cilíndrico, apresentando diversos anéis cartilaginosos unidos por cartilagens traqueais, uma espécie de tecido friboso.

Através do procedimento de traqueostomia é possível fazer uma intervenção cirúrgica caso haja alguma obstrução na traqueia, isso por meio de uma pequena abertura feita na parte frontal do pescoço introduzindo uma cânula. Nesse caso, haverá uma adaptação e adequação do indivíduo ao novo sistema como um todo, principalmente no que se refere à respiração, fala e deglutição.

2.3.2 Pulmões

O ser humano possui dois pulmões situados na cavidade do tórax, onde são promovidas as trocas gasosas no processo respiratório. É esponjoso com formato de cone irregular, sendo que o pulmão esquerdo possui dois lobos (superior e inferior) e é ligeiramente menor por conta do espaço ocupado pelo coração.

Já o direito é dividido por duas fissuras que o dividem em três lobos, o que inclui o lobo médio em relação ao anterior.

O referido órgão é recoberto pela pleura, membrana fina e transparente composta por duas camadas (pleuras parietal e visceral) que revestem os pulmões e o interior da parede torácica. Há, entre essas camadas, uma cavidade pleural que contém apenas uma fina camada de líquido que lubrifica (líquido pleural), o que permite o deslizamento suave dos pulmões na expansão e contração durante a entrada e saída de ar na produção da voz e respiração.

2.3.3 Brônquios

Na inspiração, o fluxo de ar chega até os pulmões através da traqueia, a qual se bifurca nos dois brônquios (direito e esquerdo), os quais se subdividem em ramos menores, os bronquíolos, que no final comportam pequenos sacos de ar chamados de alvéolos. É todo um processo químico-molecular que permite a absorção de parte do oxigênio do ar e a formação de gás carbônico para geração do fluxo de ar necessário à produção da voz e expiração.

Os brônquios principais fazem a ligação da traqueia aos pulmões, possuindo células produtoras de muco que absorvem partículas que adentraram, mesmo após todo processo de filtragem estabelecido no sistema respiratório. Todas essas impurezas são varridas através da movimentação dos cílios para a faringe, onde são engolidas, digeridas e eliminadas.

2.3.4 Diafragma

No que se refere à geração de fluxo de ar necessário à respiração e produção de sons, este órgão possui funcionalidade especial, uma vez que auxilia nos movimentos de inspiração e expiração, assessorado por outros diversos músculos. Do grego antigo, possui significado de divisória, uma vez que separa a cavidade abdominal da torácica, sendo o músculo de maior importância para a respiração.

O diafragma é um músculo estriado esquelético, em formato de cúpula, que se contrai quando o ar entra, juntamente com os músculos intercostais, descendo e fazendo com que as costelas subam e que haja aumento do volume da caixa torácica, o que força o ar a entrar nos pulmões. Já na expiração, o mesmo se relaxa, baixando as costelas, o que reduz o volume da caixa torácica, assim forçando o ar a sair dos pulmões.

O controle do diafragma é uma técnica muito utilizada para desenvolvimento da voz e controle do fluxo de ar que sai dos pulmões. A pressão do ar na geração da voz humana não deve ser baixa sob o risco das cordas vocais não vibrarem o suficiente, nem muito menos alta para não lesionar algum órgão do aparelho fonador.

Vale salientar que é devido a uma contração involuntária (espasmos) neste órgão que acontece o soluço, o qual pode durar desde minutos até semanas, podendo estar associado a diversas patologias, sendo mais comum em bebês por questões do sistema nervoso estar ainda imaturo. O som gerado durante a sua ocorrência é por causa do fechamento glótico repentino.

Para além dos órgãos associados ao aparelho fonador humano, ainda há algumas propriedades físicas que também são responsáveis pela geração dos fonemas, a exemplo do efeito de ressonância que define o timbre do indivíduo [16]. Durante o processo de fonação, os sinais sonoros são gerados, filtrados e amplificados.

Neste trabalho, será dada maior ênfase para a produção das vogais sustentadas, uma vez que estes sons podem ser usados para casos de patologias e para padronização de novas medidas e parâmetros [17]. A análise das diferentes características do som reproduzido, a exemplo do *jitter*, pode nos dar indícios da presença de não-conformidades no aparelho fonador do indivíduo.

Segundo [18], as vogais são sons vozeados e o mesmo afirma que, em sua produção, os articuladores orais não encontram obstáculos na passagem pela cavidade oral. O movimento do corpo da língua para estas produções se restringe a uma certa área do trato vocal, sendo que as vogais são classificadas conforme os movimentos da língua nos eixos verticais e horizontais. Assim sendo, modelar o sistema fonador por vezes torna-se uma tarefa árdua, principalmente no que se refere aos diversos aspectos associados ao sistema em questão.

Neste trabalho, serão implementados modelos matemáticos do sinal de voz, a fim de melhor representar os sons vocálicos, utilizando-se de diferentes tipos de pulsos glotais, implementando *jitter* através de processos estocásticos e fazendo uso deste fenômeno para simular vozes com e sem indícios de patologias.

Capítulo 3

Modelagem matemática determinística da produção da fala: o modelo fonte-filtro

Desde a década de 50, houve um acentuado aumento dos estudos voltados à análise da voz e da teoria que melhor explicasse a sua geração, o que implicou na criação de inúmeros laboratórios de voz com tecnologias rudimentares. Ao longo dos anos, foram sendo aprimoradas as técnicas e criados modelos com diferentes abordagens para melhor definição do sistema acústico humano, a exemplo do modelo fonte-filtro largamente utilizado nos dias atuais. Essa teoria acústica apresentada por Gunnar Fant [19], em 1970, consiste em modelar o mecanismo de produção da fala como a convolução entre uma fonte de excitação (o pulso glotal) e um sistema de filtros digitais lineares [20], representando o trato vocal e a irradiação pelos lábios/narinas, conectados em série. Dessa forma, através dos ciclos de abertura e fechamento das pregas vocais, é gerado o sinal glotal que é filtrado e amplificado pelo sistema ressonador. O modelo é apresentado na Fig. 3.1.

Figura 3.1: Modelo fonte-filtro do processo de produção da fala.

No domínio do tempo, o sinal de voz produzido pode ser representado pela convolução:

$$s[n] = g[n] * t[n] * l[n] \quad , \tag{3.1}$$

onde $s[n]$ é o sinal da fala, $g[n]$ é o sinal da fonte de excitação (sinal glotal), $t[n]$ é a resposta ao impulso do trato vocal e $l[n]$ é a resposta ao impulso da boca (irradiação).

Neste sistema apresentado, pressupõe-se que a fonte de energia e os filtros são independentes entre si, ou seja, a frequência de vibração das cordas não é alterada ao passar pelo trato vocal. É possível, assim, simular sons que se assemelham aos sons produzidos pelo aparelho fonador humano. São, então, verificadas as variáveis que estão associadas ao sistema e são analisados os diferentes padrões de sons. Vale ressaltar que os objetivos dessas análises estão relacionados com o tipo de amostra de voz que é utilizada, a exemplo de vogais sustentadas que são largamente utilizadas na identificação de padrões de patologias [17].

3.1 Fonte de excitação

Na literatura, encontramos vários modelos matemáticos determinísticos para o sinal glotal, como o modelo de Rosenberg [21], o de Fant [22], o de Liljencrants-Fant(LF) [23], e o de Klatt [24]. Uma característica importante desses modelos de pulsos glotais é a diferenciabilidade no tempo, com exceção dos momentos de abertura e fechamento da glote, sendo a fase da primeira maior que a do segundo.

Segundo [25], há vários modelos glotais que diferem em forma e complexidade tornando-os capazes de modelar diferentes estilos de vozes e, da forma mais adequada, as propriedades fisiológicas da glote. Estes podem se dividir em três categorias:

1. Modelo glotal paramétrico não interativo: Este modelo assume a não interação entre a fonte e o trato glotal, o que possibilita serem gerados pulsos através de funções paramétricas capazes de descrever o pulso glotal ou sua derivada;

2. Modelo glotal paramétrico interativo: Neste modelo há a interação entre a glote e o trato vocal, sendo assim modelados os pulsos através de funções paramétricas mais sofisticadas; e

3. Modelo glotal fisiológico: Este tipo de modelo é o mais complexo, pois envolve teorias relacionadas à interação entre a fonte glotal e as cavidades sub e supra-glotais que resultam na vibração das pregas vocais.

Tendo em vista que os diferentes modelos de pulsos glotais, no domínio temporal, não usam o mesmo número de variáveis e trabalham com nomenclaturas diferentes para parâmetros semelhantes, há uma certa problemática no que se refere à comparação dos modelos e análise de sensibilidade. Dessa forma, Natalie Henrich em [26] propôs uma estrutura unificada para estudar as propriedades dos modelos do fluxo glótico nos domínios de tempo e frequência, onde conseguiram mostrar que os diferentes modelos podem ser representados por um conjunto comum de cinco parâmetros elencados a seguir e ilustrados na Fig. 3.2.

1. Amplitude de vozeamento (A_v): O parâmetro é definido como a diferença entre o valor mínimo e máximo do fluxo glótico;

2. Velocidade de fechamento (E_e): Este parâmetro é definido como o valor mínimo da derivada do fluxo glótico e corresponde à velocidade do fluxo no instante de fechamento;

3. Quociente de abertura (O_q) : É definido como a relação entre a duração da fase aberta e o período fundamental. Em teoria, esse parâmetro pode ter valores entre 0 (sem abertura) e 1 (sem fechamento ou um fechamento incompleto). Na prática, entretanto, ele pode variar entre 0.3 e 0.98. Este parâmetro pode ser expresso como: $O_q = T_e/T_0$;

4. Coeficiente de assimetria (α_m): Constante adimensional que define o instante de máximo da onda de fluxo glótico relativo a O_q e T_0; varia entre 0.5 (forma simétrica) e 1 (caso limite de uma forma muito assimétrica). Este parâmetro pode ser expressado como $\alpha_m = T_p/(O_q T_0)$; e

5. Constante de tempo de fase de retorno (T_a): O parâmetro corresponde à diferença entre o instante do máximo de excitação e o instante onde o fluxo é quase o seu valor mínimo. O valor deste parâmetro é nulo em caso de um fechamento abrupto.

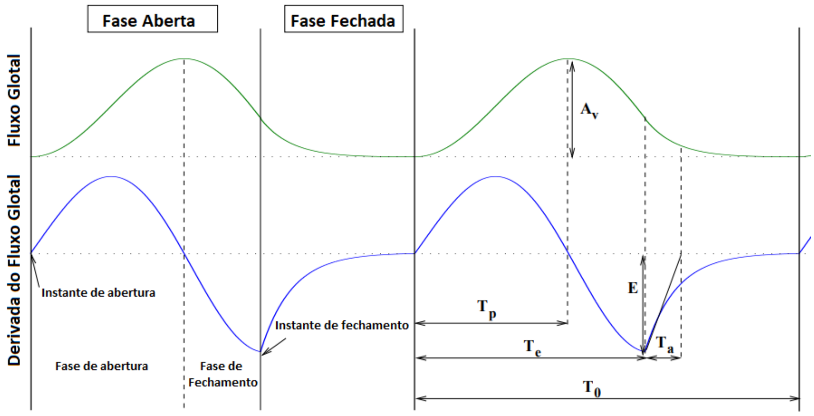


Figura 3.2: Ilustração da terminologia utilizada para descrever a fonte glótica [26].

Na Fig. 3.2, tem-se que a linha superior verde representa o sinal glotal e a linha inferior azul indica a derivada deste sinal, sendo representados ambos no domínio do tempo. No estudo aqui apresentado são utilizados dois modelos de pulsos glotais, o de Rosenberg e o de LF, ambos modelos glotais paramétricos unificados não interativos.

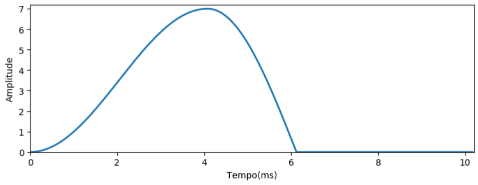
3.1.1 O modelo de Rosenberg

O modelo de Rosenberg é descrito pela Eq. (3.2) :

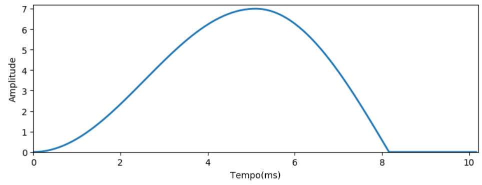
$$U_g(t) = \begin{cases} 0.5A_v[1 - cos(\pi t/T_p)] & , \ 0 \leq t \leq T_p \\ A_v cos(\pi(t - T_p)/(2T_n)) & , \ T_p < t \leq T_p + T_n \\ 0 & , \ T_p + T_n < t \leq T_0 \,. \end{cases} \tag{3.2}$$

Este modelo trigonométrico possui duas funções para representar as fases de abertura e fechamento da velocidade do volume do fluxo glótico, onde A_v é uma constante relacionada à amplitude do pulso, $T_p = \alpha_1 T_0$ é o tempo de abertura e $T_n = \alpha_2 T_0$ é o tempo de fechamento, sendo α_1 e α_2 parâmetros relacionados à porção do pulso ascendente e à porção do pulso descendente, respectivamente. Em relação aos parâmetros unificados, há uma relação direta estabelecida para este modelo e, para isso, basta tomar $T_p = \alpha_m O_q T_0$; e $T_n = T_e - T_p = (1 - \alpha_m) O_q T_0$. Sendo assim, essas equações resultam em $O_q = \alpha_1 + \alpha_2$ e $\alpha_m = \alpha_1/(\alpha_1 + \alpha_2)$. O modelo com parâmetros unificados para este modelo de pulso glotal foi proposto por Boris Doval e Christophe d'Alessandro em [27].

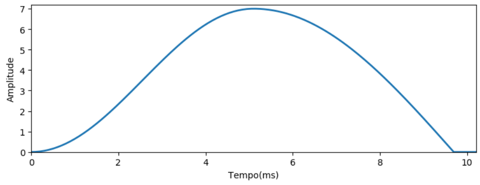
Exemplos de pulsos de Rosenberg podem ser visualizados na Fig. 3.3 .



(a) Caso 1: $\alpha_1 = 40$ e $\alpha_2 = 20$.



(b) Caso 2: $\alpha_1 = 50$ e $\alpha_2 = 30$.



(c) Caso 3: $\alpha_1 = 50$ e $\alpha_2 = 45$.

Figura 3.3: Exemplos de pulsos glotais através do modelo de Rosenberg com frequência fundamental $f_0 = 98Hz$ e $A_v = 7$.

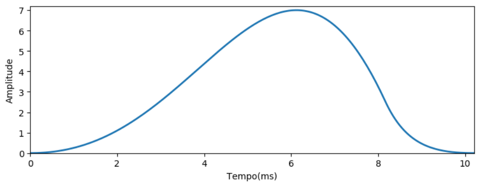
Pode ser observado que os parâmetros α_1 e α_2 controlam o tempo da fase de abertura e da fase de fechamento dos pulsos glotais. É importante ressaltar que, em casos reais, o intervalo de fechamento é menor do que o intervalo de abertura, já que as cordas vocais se fecham mais rapidamente do que se abrem.

3.1.2 O modelo de Liljencrants-Fant (LF)

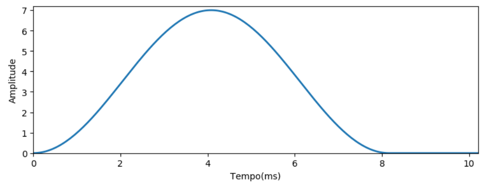
O segundo modelo, o de LF com parâmetros unificados proposto por [27], é representado pela Eq. (3.3). Este modelo é composto por uma parte senoidal modulada por uma exponencial ascendente (entre 0 e O_qT_0), seguida de uma fase de retorno exponencial decrescente (entre O_qT_0 e T_0).

$$U_g(t) = \begin{cases} -\dfrac{E_e e^{-aO_qT_0}}{sen\left(\dfrac{\pi}{\alpha_m}\right)\left(a^2 + \left(\dfrac{\pi}{\alpha_m O_q T_0}\right)^2\right)} \left(\dfrac{\pi}{\alpha_m O_q T_0} + ae^{at} sen\left(\dfrac{\pi}{\alpha_m O_q T_0} t\right) - \dfrac{\pi}{\alpha_m O_q T_0} e^{at} cos\left(\dfrac{\pi}{\alpha_m O_q T_0} t\right)\right) & , 0 \leq t < O_qT_0 \\ \\ -E_e\left(\dfrac{1}{\varepsilon Q_a(1-O_q)T_0} - 1\right)\left(T_0 - t + \dfrac{1-e^{\varepsilon(T_0-t)}}{\varepsilon}\right) & , O_qT_0 \leq t \leq T_0\,. \end{cases} \tag{3.3}$$

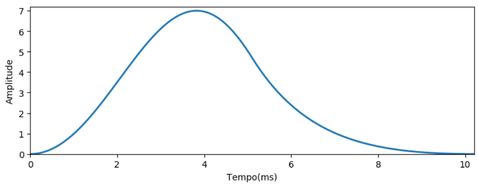
A expressão temporal da Eq. (3.3) corresponde ao modelo no caso de um fechamento não abrupto ($Q_a > 0$), cujo parâmetro Q_a tem relação com a duração efetiva da fase de retorno através de $T_a = Q_a(1 - O_q)T_0$ e é conhecido como o coeficiente da fase de retorno. O parâmetro O_q é o quociente aberto e está relacionado com o instante de fechamento glótico, como já citado. O período fundamental é T_0 e α_m é o coeficiente de assimetria, o qual, devido à fase de abertura glotal ser sempre maior do que a fase de fechamento glotal, a sua faixa de valores fica restrita entre 0.5 e 1. O parâmetro E_e representa a máxima amplitude de excitação da derivada do fluxo glótico. As constantes a e ε são obtidas através das condições de continuidade do fluxo glótico e de sua derivada no instante de excitação máxima. Exemplos de pulsos de LF podem ser visualizados na Fig. 3.4.



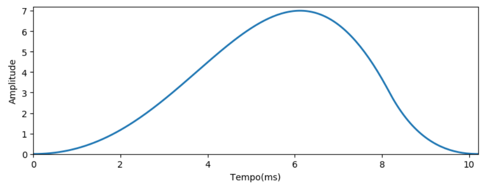
(a) Caso 1: $O_q = 0.80$, $\alpha_m = 0.75$ e $Q_a = 0.30$.



(b) Caso 2: $O_q = 0.80$, $\alpha_m = 0.50$ e $Q_a = 0.30$.



(c) Caso 3: $O_q = 0.50$, $\alpha_m = 0.75$ e $Q_a = 0.30$.



(d) Caso 4: $O_q = 0.80$, $\alpha_m = 0.75$ e $Q_a = 0.50$.

Figura 3.4: Exemplos de pulsos glotais através do modelo de LF com frequência fundamental $f_0 = 98Hz$ e $A_v = 7$.

É importante ressaltar aqui a significância que possui a modelagem do sinal glotal, uma vez que esta está intimamente ligada à qualidade do som que é reproduzido, dado que as variáveis, a exemplo da frequência fundamental, que é fator responsável pela classificação da voz como grave ou aguda, e o conjunto de harmônicos, que definem o timbre, são consequências diretas dos parâmetros distintos adotados inicialmente na simulação.

Ressalta-se, mais uma vez, a diferença entre os dois modelos apresentados. Um deles, bem mais simples, gera um fechamento abrupto das cordas vocais. O segundo, mais próximo da realidade, permite controlar a fase de fechamento, que será mais suave.

3.2 Trato vocal

O som é uma onda mecânica longitudinal que é gerada a partir das variações de pressão acústica provenientes da vibração de partículas do meio material de propagação. No caso em estudo, tem-se o trato vocal como o filtro ressonante que faz com que o fluxo de ar, que passa pelas cordas vocais, chegue aos lábios e produza a voz humana. As configurações dos nervos e músculos presentes no trato definem os formantes, moldando assim o espectro de frequência do som emitido e modificando a intensidade dos harmônicos nas frequências de ressonância.

Nesse contexto, os formantes, como frequências naturais de ressonância do trato vocal, mais especificamente na posição articulatória da vogal falada, são os principais identificadores das vogais, contribuindo de forma efetiva na construção do timbre do indivíduo. Sendo assim, em uma análise acústica, observa-se que os primeiros cinco formantes são os de maior interesse, sendo que os três primeiros são responsáveis pela identidade das vogais e possuem características instáveis, já que podem apresentar variações de vogal para vogal, enquanto que o quarto e o quinto formantes não têm a mesma variação, então sendo considerados mais estáveis, e assim responsáveis pela qualidade e brilho da voz [28].

Segundo [29], na produção de cada som, os articuladores (língua, lábios e etc.) irão se posicionar de forma determinada de maneira a obter as frequências específicas, uma vez que as frequências de ressonância de um [a], por exemplo, não são as mesmas de um [i].

Na modelagem do trato vocal, uma detalhada teoria acústica deveria levar em conta inúmeros efeitos causados pela passagem do fluxo de ar pela boca. Contudo, configurações mais simples podem ser utilizadas, a exemplo de um tubo com seção transversal variável em função do comprimento, num estado estacionário, o qual tem se mostrado satisfatório para a simulação de vogais sustentadas.

A referida modelagem do sistema possui 4 ou 5 formantes importantes, conforme ilustra a Fig. 3.5, e elas variam conforme o trato vocal vai dilatando e contraindo durante a pronúncia dos sons diversos, fazendo com que os harmônicos do pulso glotal sejam filtrados e ampliados. A função de transferência do trato vocal na produção de vogais orais se caracteriza por ter apenas pólos, que correspondem aos formantes da fala vozeada. As larguras de banda são determinadas nos pontos de média potência (3 dB abaixo) da curva de resposta em frequência. Estas últimas são fatores que consideram os efeitos de perdas por paredes moles, fricção e condução térmica no sistema.

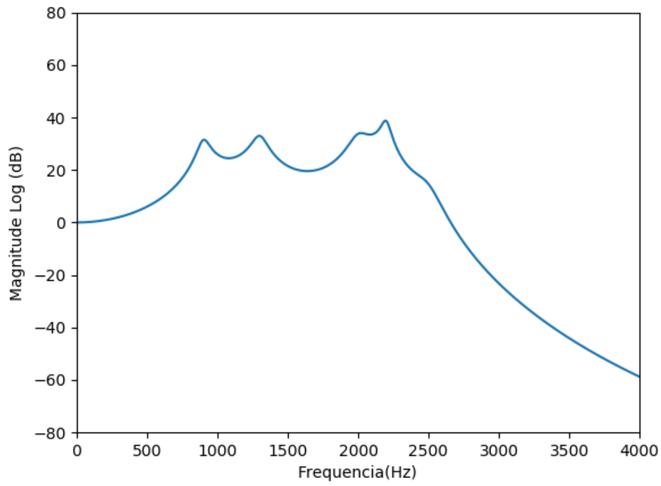


Figura 3.5: Exemplo de resposta em frequência do trato vocal para os formantes $F_1 = 900$, $F_2 = 1300$, $F_3 = 2000$, $F_4 = 2200$ e $F_5 = 2500$ [30].

Durante a emissão sonora, a energia dos vários harmônicos da fonte glotal não é transmitida igualmente, uma vez que as frequências baixas, que incluem aí também as menores formantes, concentram maior energia fornecida pela fonte glotal. Dito isto, a utilização dos 3 primeiros formantes em muitos casos já é suficiente para lidar com a maioria das variações fonéticas para as vogais. Vale salientar que, embora a largura de banda do formante não seja necessariamente um fator crítico na percepção de vogais, há possivelmente uma largura de banda ótima que facilita a discriminação e identificação de vogais [31].

Vale salientar que a produção de fala exige boas condições das estruturas do trato vocal, uma vez que qualquer interferência na dinâmica implica na geração de um sinal acústico diferente do que a pessoa emite normalmente, geralmente exigindo algum esforço adicional da mesma na sua produção [25].

3.3 Radiação labial

A radiação é uma terminologia que se refere à forma como o trato vocal termina na atmosfera [31]. Está relacionado ao efeito de filtragem que ocorre quando o som sai da boca e é espalhado no espaço em todas as direções. Assume-se que o som de saída aumenta em frequência a uma taxa de 6 dB/oitava, comportando-se como um diferenciador, ou seja, multiplicando (ou somando em dB) o sinal por uma reta de declive +6dB/oitava no domínio da frequência.

A radiação labial pode ser aproximada através de um filtro FIR (resposta ao impulso finita) passa-alta de primeira ordem, tendo como função de transferência $R(z) = 1 - \alpha z^{-1}$, para $\alpha \in [0.95, 0.99]$. Seu efeito é ampliar as componentes de alta frequência do sinal referente à cavidade bucal.

Capítulo 4

O fenômeno do *jitter*

A frequência fundamental da voz está intimamente ligada às características do sistema fonador humano, em especial à fisiologia das cordas vocais, o que implica nas diferenças entre os sons emitidos por homens e mulheres, conforme já foi relatado. Como as pregas vocais não são exatamente simétricas e não abrem e fecham com total sincronismo, o fenômeno denominado jitter ocorre no sistema. É um fenômeno presente em todas as vozes humanas e provoca pequenas perturbações aleatórias na frequência fundamental, como pode ser notado na Fig. 4.1. É uma característica acústica do sinal de voz que é largamente utilizada para detecção de patologias no aparelho fonador humano.

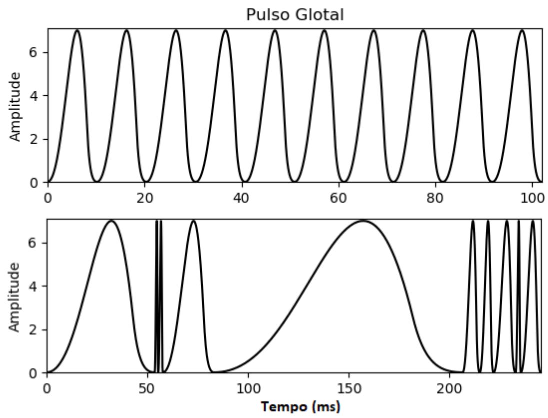


Figura 4.1: Exemplos de trens de pulsos glotais sem *jitter* (acima) e com *jitter* exagerado (abaixo), respectivamente.

Para a identificação dos parâmetros da voz, atualmente são utilizadas diversas ferramentas computacionais, a exemplo dos programas Praat e o MDVP (*Multi-Dimensional Voice Program*) [32], os quais são *softwares* de análise acústica. Dessa forma, para facilitar comparação e a generalização dos resultados do estudos usando diferentes programas, faz-se necessário o conhecimento das diferenças entre os programas e suas medidas, até mesmo para escolher qual o programa mais indicado para cada situação [33].

A medição do *jitter* presente na fala humana tem se tornado cada vez mais um assunto de relevada significância para o meio do processamento digital de voz, principalmente pela vantagem de permitir ser feita uma análise do aparelho fonador não-invasiva, objetiva e de baixo custo de implementação, no que tange à existência de patologias no trato vocal do indivíduo.

4.1 Cálculo do *jitter*

Para determinação do *jitter*, são utilizados diferentes tipos de medidas para melhor representar o fenômeno, os quais são listados abaixo, baseados na análise de voz utilizando o *software* Praat [32]. Vale salientar que nas Eqs. (4.1), (4.2), (4.3), (4.4) e (4.5), tem-se T_i como o intervalo de tempo do pulso i, $i \in \{1, 2, 3, ..., N\}$, N como o número de pulsos executados e Jit_{abs}, Jit_{loc}, Jit_{rap}, Jit_{ppq5} e Jit_{ddp} como as medidas de dispersão do *jitter*.

4.1.1 *Jitter* absoluto

Representa a média das diferenças absolutas entre dois períodos consecutivos do pulso glotal identificado no sistema, sendo conhecido como *Jita*.

$$Jit_{abs} = \frac{1}{N-1} \sum_{i=1}^{N-1} |T_i - T_{i+1}| \,. \tag{4.1}$$

4.1.2 *Jitter* local

Representa a diferença média absoluta entre dois períodos consecutivos, dividida pelo período médio, sendo conhecido como *Jitt*. O MDVP considera que, para valores acima de 1,040%, há um indício da existência de patologia no sistema fonador do indivíduo.

$$Jit_{loc} = 100 \left(\frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_i - T_{i+1}|}{\frac{1}{N} \sum_{i=1}^{N} T_i} \right) . \tag{4.2}$$

É possível observar na Fig. 4.1 uma demonstração do efeito *jitter* no pulso glotal gerado. Apesar de no exemplo haver um valor de *jitter* elevado, com $Jit_{loc} = 101.4\%$, este foi apenas para que fosse possível a visualização do fenômeno em questão.

4.1.3 *Jitter* RAP

A variável em questão, RAP (*Relative Average Perturbation*), representa a diferença média absoluta entre um período e a média desse mesmo período com seus dois vizinhos, dividida pelo período médio. A ferramenta computacional MDVP considera que, para valores acima de 0,680%, há um indício da existência de patologia no sistema fonador do indivíduo.

$$Jit_{rap} = 100 \left(\frac{\frac{1}{N-2} \sum_{i=2}^{N-1} |T_i - \frac{1}{3}(T_{i-1} + T_i + T_{i+1})|}{\frac{1}{N} \sum_{i=1}^{N} T_i} \right) . \tag{4.3}$$

4.1.4 *Jitter* PPQ5

A variável em questão, PPQ5 (*Five Point Period Perturbation Quotient*), representa a diferença média absoluta entre um período e a média desse mesmo período com seus quatro vizinhos próximos, dividida pelo período médio. A ferramenta computacional MDVP considera que, para valores acima de 0,840%, há um indício da existência de patologia no sistema fonador do indivíduo.

$$Jit_{ppq5} = 100 \left(\frac{\frac{1}{N-4} \sum_{i=3}^{N-2} |T_i - \frac{1}{5}(T_{i-2} + T_{i-1} + T_i + T_{i+1} + T_{i+2})|}{\frac{1}{N} \sum_{i=1}^{N} T_i} \right) . \tag{4.4}$$

4.1.5 ***Jitter*** **DDP**

DDP (*Difference of Differences of Periods*) representa a diferença média absoluta das consecutivas diferenças entre períodos vizinhos, dividida pelo período médio.

$$Jit_{ddp} = 100 \left(\frac{\frac{1}{N-2} \sum_{i=2}^{N-1} |(T_{i+1} - T_i) - (T_i - T_{i-1})|}{\frac{1}{N} \sum_{i=1}^{N} T_i} \right) . \tag{4.5}$$

Capítulo 5

Modelos estocásticos do sinal glotal

Em síntese de voz, há a necessidade de que sejam criados modelos que contenham cada vez mais os fenômenos relacionados à fala natural. Para o presente estudo, optou-se por representar o sistema fonador humano através do modelo fonte-filtro, o qual tem se mostrado bastante eficiente na geração da voz. Contudo, os mais reais apresentam incertezas, algumas das quais podem ser consideradas aleatórias e, assim, há uma recorrente tentativa de caracterizar estas através de processos estocásticos. Neste trabalho, modelos estocásticos para o pulso glotal são construídos a fim de simular o efeito *jitter* e, assim, tornar os modelos mais próximos da realidade.

Na literatura, há diversas formas de simular o *jitter* gerado pela voz humana. Em [2], por exemplo, utiliza-se um sistema massa-mola-amortecedor para representar a dinâmica das cordas vocais que, através de um sinal de voz obtido experimentalmente, são identificados os parâmetros ótimos do modelo e definidas as funções densidades de probabilidade da variável aleatória relacionada à frequência fundamental. Através da construção do processo estocástico associado à rigidez elástica da mola do modelo, foi possível controlar as variações do período fundamental em torno do seu valor médio, assim configurando a implementação do *jitter*.

Em [3], através do modelo de fita [34] para a oscilação das cordas vocais, o autor considera a perturbação da frequência fundamental instantânea e microtremores vocais para a construção de cinco modelos estocásticos de *jitter*. É gerado um processo estocástico para simulação das pequenas flutuações, em torno do valor médio do período fundamental da voz sintetizada, e são feitas análises no que tange à sensibilidade das variáveis utilizadas no sistema.

É importante ressaltar a grande importância dos modelos citados, principalmente no que diz respeito às suas diferentes abordagens, que têm como objetivo a construção de sons mais naturais com *jitter*. No presente trabalho, o intervalo glotal é considerado um processo estocástico e, junto com a sua modelagem, caracterizam ineditismo na formação do *jitter* presente no sinal glotal. Dois modelos distintos de densidade espectral são associadas ao processo estocástico que modela o intervalo glotal. Os modelos determinísticos de sinal glotal considerados são o modelo de Rosenberg [21] e o modelo de LF com parâmetros unificados [35]. Para cada um desses modelos, o intervalo glotal é considerado como um processo estocástico. Para cada modelo, duas densidades espectrais distintas são associadas e um estudo comparativo é realizado. Primeiro, será apresentado o modelo do processo estocástico associado ao intervalo glotal. A partir daí, quatro casos serão considerados:

1. Modelo de pulso glotal de Rosenberg e densidade espectral a 2 parâmetros;

2. Modelo de pulso glotal de Rosenberg e densidade espectral a 3 parâmetros;

3. Modelo de pulso glotal de LF e densidade espectral a 2 parâmetros; e

4. Modelo de pulso glotal de LF e densidade espectral a 3 parâmetros .

5.1 Modelagem estocástica do intervalo de tempo glotal

O intervalo de tempo glotal, isto é, o intervalo de tempo necessário para que as cordas vocais se abram e se fechem completamente, será considerado um processo estocástico e sua construção será explicada a seguir.

Seja t_g o intervalo de tempo glotal, isto é, o intervalo de tempo associado a um pulso glotal, no caso determinístico. No caso real, t_g terá comprimento variável, isto é, para cada ciclo glotal, t_g terá um valor diferente, sendo considerada uma variável aleatória denotada por T_g.

No caso determinístico, ao discretizarmos o sinal, o intervalo de tempo glotal pode ser dividido em N intervalos; isto é, $t_g = N\delta t = \sum_{i=1}^{N} \delta t$, sendo δt fixo. No caso aleatório, teremos $T_g = \sum_{i=1}^{N} \Delta t(t_i)$, sendo $\Delta t(t)$ um processo estocástico que será descrito adiante.

Consideremos T_{g_i} cada instante glotal, com $i = 1, \dots, N$. Assim,

$$T_{g_{i+1}} = T_{g_i} + \Delta t(t_i), \quad i = 1, ..., N. \tag{5.1}$$

Tem-se que $\Delta t(t_i)$ são amostras de uma realização do processo estocástico $\Delta t = \{\Delta t(t), t \in \mathbb{R}\}$ e que, a partir do modelo estocástico proposto por [2], o processo estocástico $\Delta t(t)$ foi construído, com as seguintes características:

1. Para todo t, $0 < \Delta t_0 \leq \Delta t(t)$, onde Δt_0 é uma constante positiva independente de t.

2. É um processo estacionário, não gaussiano, uma vez que só pode ter valores positivos.

3. $E\{(\Delta t(t))^2\} < +\infty$ para todo t (processo estocástico de segunda ordem), tal que $E\{\Delta t(t)\} = \underline{\Delta t(t)} > \Delta t_0 > 0$, sendo considerado contínuo em média-quadrática para garantir a existência de uma unidade espectral de potência.

Aqui, intoduziremos um processo estocástico real gaussiano de segunda ordem, $Y = \{Y(t), t \in \mathbb{R}\}$, centrado e contínuo em média quadrática, estacionário e ergódico, fisicamente realizável. A representação do processo estocástico $\Delta t(t)$ pode ser escrita como:

$$\Delta t(t) = \Delta t_0 + (\underline{\Delta t} - \Delta t_0)(\underline{y} + Y(t))^2, \quad \forall t \in \mathbb{R}, \tag{5.2}$$

onde $\underline{y}$ é um parâmetro tal que:

$$E\{(\underline{y}+Y(t))^2\} = 1 \text{ e } E\{(\underline{y}+Y(t))^4\} < +\infty\,, \tag{5.3}$$

uma vez que $E\{\Delta t(t)\} = \underline{\Delta t}$ e $E\{(\Delta t(t))^2\} < +\infty$.

O processo estocástico gaussiano $Y(t)$ é construído como um filtro linear, $Y = h * N_\infty$, dado pela convolução do ruído branco gaussiano centrado N_∞, cuja função densidade espectral de potência é constante e igual a $S_{N_\infty}(\omega) = 1/(2\pi)$ para todo real w, pelo filtro $h = \mathfrak{F}^{-1}\{H\}$, que é a transformada inversa de Fourier da função resposta em frequência $H(\omega)$.

Dois casos de $H(\omega)$ serão considerados nesse texto:

1. $H(\omega) = \dfrac{a}{i\omega+b}$, dependendo de 2 parâmetros: a e b; e
2. $H(\omega) = \frac{a}{-\omega^2+2i\omega\xi b+b^2}$, dependendo de 3 parâmetros: a, b e ξ.

Assim, duas densidades espectrais de potência de Y serão consideradas, respectivamente:

$$S_Y(\omega) = \frac{1}{2\pi}\frac{a^2}{\omega^2+b^2} \quad , \quad a > 0 \quad e \quad b > 0; \quad e \tag{5.4}$$

$$S_Y(\omega) = \frac{1}{2\pi}\frac{a^2}{(b^2-\omega^2)^2+4\xi^2b^2\omega^2} \quad , \quad a > 0 \quad , \quad b > 0 \quad e \quad \xi > 0. \tag{5.5}$$

Pode ser deduzido que o processo diferenciável $\{\dot{Y}(t), t \in \mathbb{R}\}$ do processo estocástico $\{Y(t), t \in \mathbb{R}\}$ é um processo estocástico de segunda ordem, porque $\int_{\mathbb{R}}\omega^2 S_Y(\omega)d\omega < +\infty$. Para o primeiro caso, isto é, $S_Y(\omega) = \dfrac{1}{2\pi}\dfrac{a^2}{\omega^2+b^2}$, o processo $Y(t)$ é definido a partir da solução da equação diferencial estocástica de Itô:

$$dZ(t) = -bZ(t)dt + adW(t), \quad t > 0. \tag{5.6}$$

Dada a condição inicial $Z(0) = 0$ a.s., onde W é processo de Wiener indexado por $[0, +\infty[$, pode ser provado [36] que a Eq. (5.6) tem uma única solução $\{Z(t), t \geq 0\}$ tal que para $t_0 \to +\infty$, o processo estocástico $\{Z(t), t \geq t_0\}$ é estocasticamente equivalente ao processo estocástico estacionário $Y(t)$, ou seja, se t_0 é escolhido suficientemente grande, $Z(t)$ e $Y(t)$ são o mesmo processo gaussiano definido anteriormente, sendo a densidade espectral de potência dada pela equação (5.4). A solução dessa equação diferencial neste trabalho foi realizada através de métodos numéricos, de forma recursiva, sendo utilizado os valores de Y depois de uma quantidade suficiente de iterações.

A solução da Eq. (5.6) foi dada através da seguinte formulação, utilizando-se do método de Euler semi-implícito:

$$Y(n+1) - Y(n) = -b\frac{Y(n)+Y(n+1)}{2}\Delta t + aN(0,1)\sqrt{\Delta t}; \quad e \tag{5.7}$$

$$Y(n+1) = \frac{(1-\dfrac{b\Delta t}{2})Y(n) + aN(0,1)\sqrt{\Delta t}}{1+\dfrac{b\Delta t}{2}}, \tag{5.8}$$

onde N(0, 1) é a distribuição normal com média zero e variância 1. As Eqs. (5.7) e (5.8) mostram como podem ser determinados os valores estocásticos de $Y(t)$ e de $Z(t)$, isto depois de considerável número de iterações.

Na Equação (5.4), a e b devem satisfazer a igualdade $E\left\{(\underline{y}+Y(t))^2\right\} = 1$, e levando em consideração que $Y(t)$ possui valor esperado nulo e $\underline{y}$ é uma constante, tem-se que:

$$\underline{y}^2 + E\{Y^2(t)\} = 1 \quad \Longrightarrow \quad \underline{y}^2 + \int_{-\infty}^{+\infty} S_Y(\omega)d\omega = 1 \quad \Longrightarrow \quad \underline{y}^2 + \int_{-\infty}^{+\infty} \frac{a^2}{2\pi(\omega^2+b^2)}d\omega = 1\,, \tag{5.9}$$

o que resulta em:

$$\underline{y}^2 = 1 - \frac{a^2}{2b}\,, \tag{5.10}$$

com as seguintes condições para a e b, uma vez que $\underline{y}^2 > 0$:

$$0 < a < \sqrt{2b} \quad , \quad b > 0. \tag{5.11}$$

A fim de garantir a convergência da Equação Diferencial de Itô para o processo estocástico de $\Delta t(t)$, algumas considerações são necessárias.

O referido método de Euler semi-implícito foi utilizado, mas é necessário que seja analisado o comportamento do seu valor esperado e do seu momento de segunda ordem, os quais são definidos a seguir nas Eq. (5.12) e (5.13), para verificação da ergodicidade e convergência:

$$E\{\Delta t(n)\} = \lim_{n\to+\infty} \overline{\Delta t}(n) \quad , \tag{5.12}$$

sendo $\overline{\Delta t}(n) = \frac{1}{n}\sum_{i=1}^{n} \overline{\Delta t}(i)$; e

$$E\{[\Delta t(n)]^2\} = \lim_{n\to+\infty} [\overline{\Delta t}(n)]^2 \quad , \tag{5.13}$$

sendo $[\overline{\Delta t}(n)]^2 = \frac{1}{n}\sum_{i=1}^{n}[\overline{\Delta t}(i)]^2$.

Para o segundo caso, isto é, $S_Y(\omega) = \dfrac{1}{2\pi}\dfrac{a^2}{(b^2-\omega^2)^2 + 4\xi^2 b^2\omega^2}$, estabelecendo-se $\{\mathbb{Z}(t) = (Y(t), \dot{Y}(t), t \geq 0)\}$ como processo estocástico com valores em $\mathbb{R}^2$, o processo $Y(t)$ é definido a partir da solução da equação diferencial estocástica de Itô:

$$d\mathbb{Z}(t) = -[\alpha]\mathbb{Z}(t)dt + [\beta]d\mathbb{W}(t), \quad t > 0, \tag{5.14}$$

com a condição inicial $\mathbb{Z}(0) = (0,0)$ a.s., onde $\mathbb{W}(t), t \geq 0$ é processo estocástico normalizado de Wiener indexado por $[0,+\infty[$, $[\alpha]$ é uma matriz real de dimensão (2x2) e $[\beta]$ é um vetor real, tal como se segue:

$$[\alpha] = \begin{bmatrix} 0 & 1 \\ -b^2 & -2\xi b \end{bmatrix} \quad , \quad [\beta] = \begin{bmatrix} 0 \\ a \end{bmatrix} \quad . \tag{5.15}$$

Pode ser provado [36] que a Eq. (5.14) tem uma única solução $\{\mathbb{Z}(t), t \geq 0\}$ tal que, para $t_0 \to +\infty$, o processo estocástico $\{\mathbb{Z}(t), t \geq t_0\}$ é estocasticamente equivalente ao processo estocástico estacionário $\{(Y(t), \dot{Y}(t)), t \in \mathbb{R}\}$.

A condição definida pela Eq. (5.3) para este caso, levando em consideração a análise feita na Eq. (5.9), implica em:

$$\underline{y}^2 + \int_{-\infty}^{+\infty} S_Y(\omega)d\omega = 1 \quad \Longrightarrow \quad \underline{y}^2 = 1 - \frac{a^2}{4\xi b^3} \quad , \tag{5.16}$$

com as seguintes condições para a, b e ξ, uma vez que $\underline{y}^2 > 0$:

$$0 < a^2 < 4\xi b^3 \quad , \quad b > 0 \quad , \quad \xi > 0 \quad . \tag{5.17}$$

A solução da Eq. (5.14) foi dada através do seguinte método numérico, utilizando-se do método de Euler semi-implícito e de resolução de equações de estado, conforme é apresentado a seguir.

$$\mathbb{Z}(n+1) - \mathbb{Z}(n) = -[\alpha]_{2x2}\frac{\mathbb{Z}(n) + \mathbb{Z}(n+1)}{2}\Delta t + [\beta]_{2x1}N(0,1)\sqrt{\Delta t}, \tag{5.18}$$

$$\begin{bmatrix} Y(n+1) - Y(n) \\ \dot{Y}(n+1) - \dot{Y}(n) \end{bmatrix} = -[\alpha]_{2x2}\begin{bmatrix} Y(n+1) + Y(n) \\ \dot{Y}(n+1) + \dot{Y}(n) \end{bmatrix}\frac{\Delta t}{2} + [\beta]_{2x1}N(0,1)\sqrt{\Delta t}, \tag{5.19}$$

$$\begin{bmatrix} Y(n+1) \\ \dot{Y}(n+1) \end{bmatrix} + [\alpha]_{2x2}\begin{bmatrix} Y(n+1) \\ \dot{Y}(n+1) \end{bmatrix}\frac{\Delta t}{2} = \begin{bmatrix} Y(n) \\ \dot{Y}(n) \end{bmatrix} - [\alpha]_{2x2}\begin{bmatrix} Y(n) \\ \dot{Y}(n) \end{bmatrix}\frac{\Delta t}{2} + [\beta]_{2x1}N(0,1)\sqrt{\Delta t}, \tag{5.20}$$

$$\left[\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{\Delta t}{2}[\alpha]_{2x2}\right]\begin{bmatrix} Y(n+1) \\ \dot{Y}(n+1) \end{bmatrix} = \left[\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \frac{\Delta t}{2}[\alpha]_{2x2}\right]\begin{bmatrix} Y(n) \\ \dot{Y}(n) \end{bmatrix} + [\beta]_{2x1}N(0,1)\sqrt{\Delta t}, \tag{5.21}$$

$$\begin{bmatrix} Y(n+1) \\ \dot{Y}(n+1) \end{bmatrix} = \left[\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{\Delta t}{2}[\alpha]_{2x2}\right]^{-1}\left[\left[\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \frac{\Delta t}{2}[\alpha]_{2x2}\right]\begin{bmatrix} Y(n) \\ \dot{Y}(n) \end{bmatrix} + [\beta]_{2x1}N(0,1)\sqrt{\Delta t}\right], \tag{5.22}$$

Após uma quantidade suficiente de iterações, através da Eq. 5.22, é possível obter os valores de $\{\mathbb{Z}(t) = (Y(t), \dot{Y}(t), t \geq 0)\}$ e verificar que o processo estocástico Y apresenta comportamento semelhante, possuindo, assim, a referida densidade espectral.

Para este modelo de processo estocástico com três parâmetros a, b e ξ, temos uma densidade espectral de potência S_Y variável que permite a simulação de diferentes padrões de *jitter* locais, calculados pela Eq. (4.2). Antes de mostrar os resultados pertinentes ao modelo, é importante estudar a expressão de S_Y a fim de mostrar como as variáveis influenciam, enfatizando, principalmente, em como se comporta o sistema com a variação de ξ.

Tomando como o valor máximo referência de $S_Y(2\pi f; a, b, \xi)_{max} = a^2/(8\pi b^4\xi^2(1-\xi^2))$, que acontece quando:

$$\omega = 2\pi f_p\sqrt{1-2\xi^2} \quad , \tag{5.23}$$

será apresentado, a seguir, uma ilustração de como ξ afeta significativamente na densidade espectral de potência. Para efeito de comparação, os valores de a, b e ξ serão fixados em 10, 360π e 0.1, respectivamente.

A Fig. 5.1 representa a função densidade espectral de potência do processo $Y(t)$ a três parâmetros normalizada, mostrando como o valor de ξ afeta significativamente na largura de banda do sistema. O valor adotado de $b = 2\pi f_p$ toma como parâmetro f_p, o qual deve ser escolhido próximo aos valores de frequências fundamentais assumidas pela voz que vier a ser simulada.

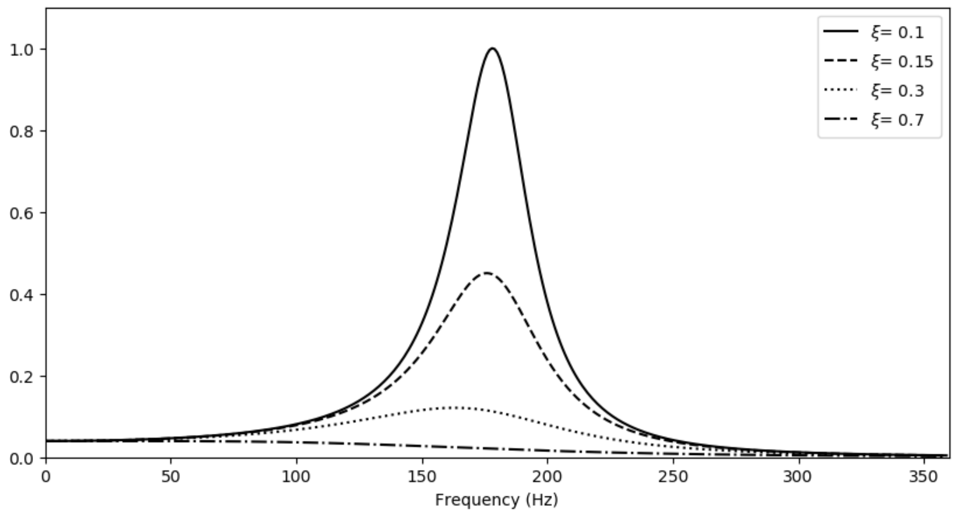


Figura 5.1: Gráfico da função $f \mapsto S_Y(2\pi f; a, b, \xi)/S_{Ymax}$, com $a = 10$, $b = 360\pi$ e $\xi \in \{0.1, 0.15, 0.3, 0.7\}$.

Tomando $b = 2\pi c_b f_p$, assumindo $f_p = 180Hz$ e $c_b \in \{1.0, 1.2, 1.5, 2.0\}$, para $a = 10$ e $\xi = 0.1$, tem-se que o ponto de máximo das curvas muda de posição, uma vez que a abscissa do ponto citado depende diretamente do valor de b, e como b está aumentando, a curva desloca para a direita, conforme indicado na Fig. 5.2. Portanto, é possível modificar a frequência da voz apenas alterando c_b através da frequência fundamental que é prefixada.

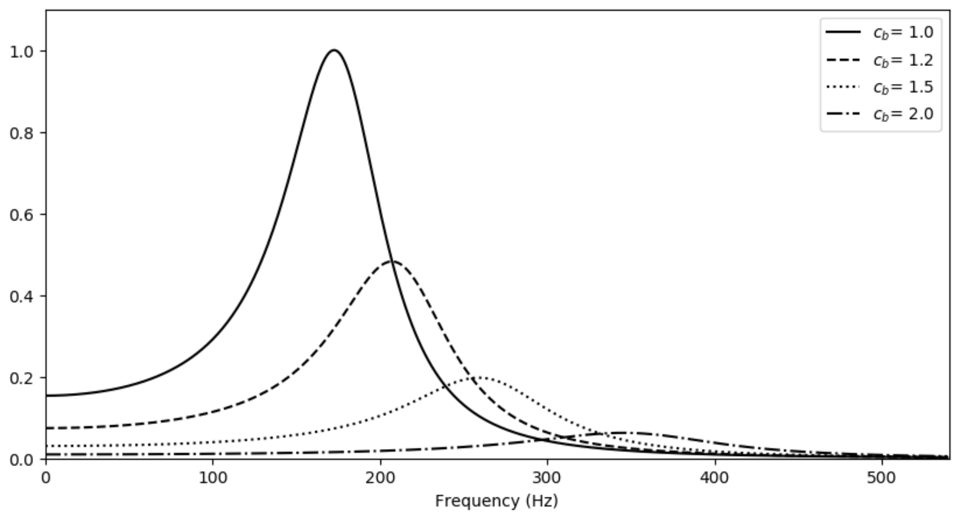


Figura 5.2: Gráfico da função $f \mapsto S_Y(2\pi f; a, b, \xi)/S_{Ymax}$, com $a = 10$, $\xi = 0.1$ e $b = 2\pi c_b f_p$, sendo $f_p = 180$ e $c_b \in \{1.0, 1.2, 1.5, 2.0\}$.

É válido citar aqui que este último modelo apresentado, além de permitir controlar a intensidade do *jitter*, consegue alterar a frequência fundamental f_0 através da manipulação das variáveis que constituem a densidade espectral de potência do sinal formado. Assim, através dos diferentes tipos de pulsos glotais utilizados, é possível criar diferentes *jitter* e alterar o colorido frequencial do sistema.

Capítulo 6

Simulações

Para a apresentação dos resultados, simulações serão realizadas a partir dos modelos construídos. Os sinais glotais serão gerados a partir de quatro casos considerados. Para cada um dos casos, sinais de vozes serão obtidos considerando combinações de parâmetros diferentes.

Os sinais de vozes serão obtidos através da teoria fonte-filtro e os parâmetros utilizados serão os seguintes:

1. Para a fonte glotal:

   Os pulsos glotais gerados com o modelo de Rosenberg apresentam $f_0 = 98Hz$, $Av = 7$, $\alpha_1 = 50$ e $\alpha_2 = 30$. Já pelo modelo de LF, foram utilizadas as seguintes constantes: $f_0 = 98Hz$, $Av = 7$, $O_q = 0.80$, $\alpha_m = 0.75$ e $Q_a = 0.30$. Foi escolhida uma frequência de amostragem para o sistema igual a $f_m = 20000Hz$, $\Delta t_0 = 1/40000$ e estipulado um tempo $T_{fs} = 3s$ de sustentação da vogal.

2. Para o trato vocal:

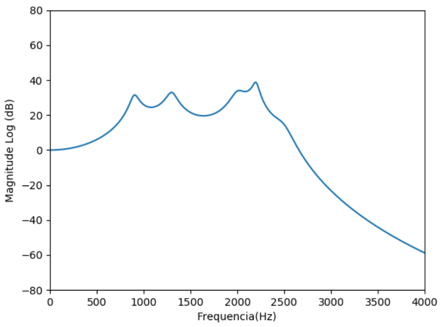
   Para o nosso estudo em questão, após a síntese dos pulsos glotais com parâmetros unificados, foram estabelecidos os parâmetros do trato vocal. Os formantes, assim como as suas respectivas larguras de banda que representam as perdas, estão definidos na Tab. 6.1. Os três primeiros formantes (F_1, F_2 e F_3) têm menor dependência com o locutor e prestam-se, principalmente, para diferenciar as vogais.

Fonemas	F_1	F_2	F_3	F_4	F_5
\a\	900	1300	2000	2200	2500
\e\	450	1700	2000	2200	2310
\i\	300	1900	2100	2200	2490
\o\	500	800	2150	2200	2490
\u\	360	700	2170	2200	2330
Banda de Passagem	41	52	70	32	100

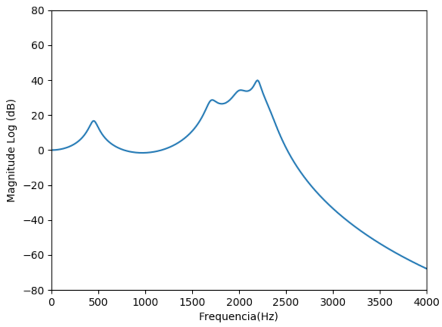
Tabela 6.1: Formantes e Bandas de Passagem para um indivíduo com classificação vocal baixo.

Os formantes superiores (F_4 e F_5), por outro lado, têm menor conteúdo linguístico e maior variação com o locutor. Acusticamente, os formantes amplificam seletivamente os harmônicos gerados pela vibração laríngea [37].

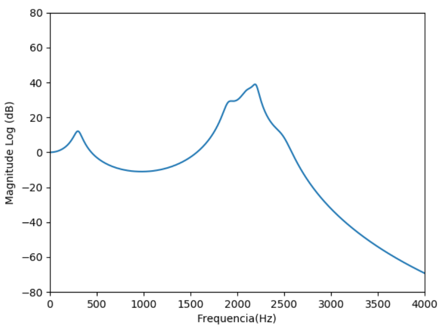
Para os formantes escolhidos, a Fig. 6.1 representa a resposta em frequência do filtro que modela o trato vocal. Eles, juntamente com as bandas de passagem, foram selecionadas da carta dos formantes do espanhol como aquelas que permitiram perceber maior clareza às vogais [30]. É possível identificar, no gráfico, a maioria dos formantes que compõem o aparelho ressonador de um indivíduo com classificação vocal baixo (voz mais grave).



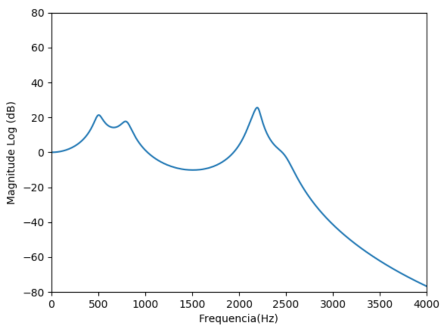
(a) Vogal a



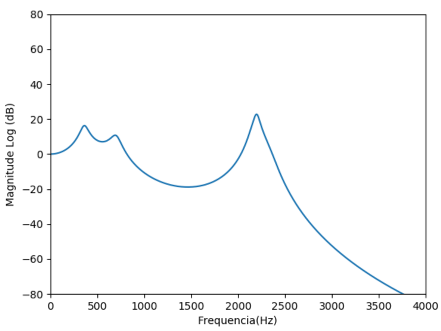
(b) Vogal e



(c) Vogal i



(d) Vogal o



(e) Vogal u

Figura 6.1: Resposta em frequência do trato vocal indicando os formantes das vogais para um indivíduo com classificação vocal baixo.

3. Para a radiação labial:

   Para implementação do filtro FIR de primeira ordem, foi escolhido o valor de $\alpha = 0.95$, o que resulta na função de transferência $R(z)$ apresentada na Eq. (6.1).

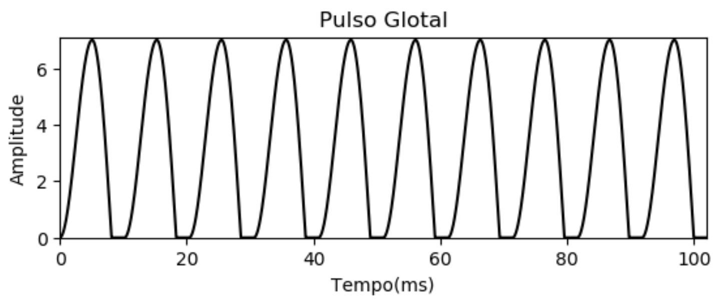
$$R(z) = 1 - 0.95z^{-1} \tag{6.1}$$

   Este filtro de pré-ênfase do sinal de voz relaciona o sinal logo depois do trato vocal $V[n]$ ao sinal irradiado pela boca $\overline{V}[n]$ através da equação diferença $\overline{V}[n] = V[n] - 0.95V[n-1]$.

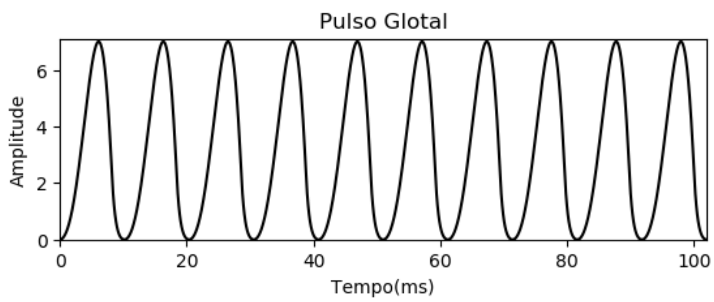
Nas simulações desenvolvidas neste trabalho, os parâmetros supracitados serão fixos e, portanto, será analisada a influência das variáveis associadas às duas diferentes densidades espectrais de potência utilizadas para os processos estocásticos estabelecidos.

Aqui, vale salientar que o processo estocástico $\Delta t(t)$ é o que constrói, iteração após iteração, o período fundamental T_g de cada ciclo, sendo este uma variável aleatória que se modifica ciclo após ciclo, contraindo ou expandindo.

Antes de inserir o processo estocástico no sistema, é importante expor os pulsos glotais sem interferência do fenômeno *jitter*. Para isso, na Fig. 6.2 temos a representação gráfica no tempo dos trens de pulsos glotais sem *jitter* para $f_0 = 98Hz$. Eles representam o caso determinístico definido pelas Eq. (3.2) e (3.3) dos pulsos glotais de Rosenberg e LF, respectivamente.



(a) Trem de pulsos de Rosenberg sem *jitter*.



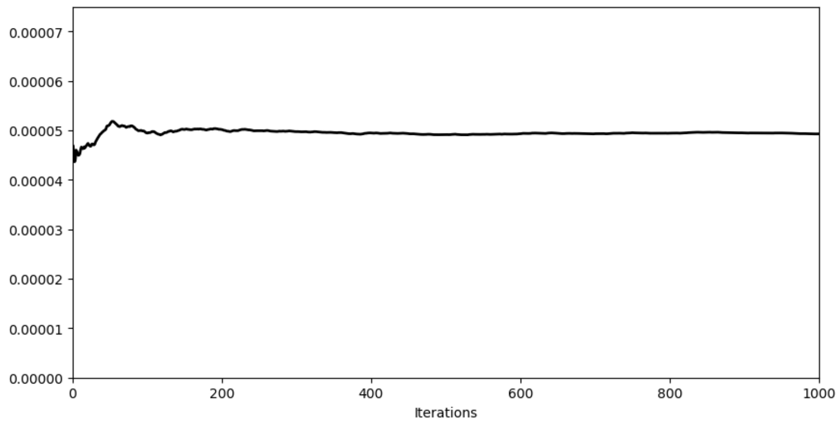
(b) Trem de pulsos de LF sem *jitter*.

Figura 6.2: Trens de pulsos sem *jitter*.

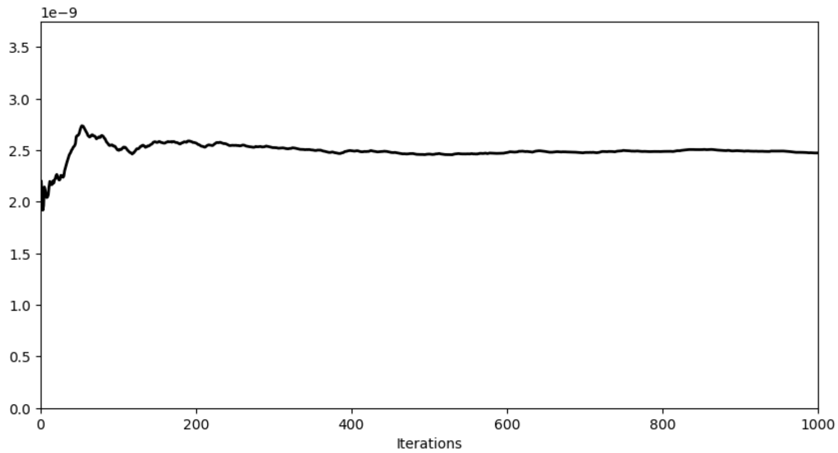
A partir daqui, serão feitas as simulações dos modelos estocásticos definidos no capítulo anterior, e para cada um destes serão verificados os resultados obtidos com a síntese dos pulsos de Rosenberg e de LF. Assim, através da linguagem de programação *Python*, serão analisados os quatro casos e construídos o sistema fonte-filtro, a excitação glotal de Rosenberg e LF com parâmetros unificados e os filtros simulando o trato vocal, suas perdas e a radiação labial. Para execução das simulações, foi utilizado um notebook Intel Core i5-8250U com memória RAM de 8GB, através do Python versão 2.7.15, cujo tempo de processamento médio das vogais, com tempo de sustentação de 3 segundos, e geração dos gráficos correspondentes duraram, em média, de 3 a 9 minutos de acordo com a complexidade do modelo.

6.1 Caso I: Modelo de Rosenberg e densidade espectral a 2 parâmetros

Para este modelo de densidade espectral, como foi dito antes, é necessário fazer análise de convergência do processo, a fim de estimar a partir de qual número de iterações o sistema passa a convergir. Nesse intuito, foi simulado o comportamento do sistema através das Eq. (5.12) e (5.13), iteração após iteração.



(a) $E\{\Delta T(n)\}$



(b) $E\{(\Delta T(n))^2\}$

Figura 6.3: Valor esperado de primeira (figura acima) e segunda (figura abaixo) ordem da variável $\Delta T(n)$, assumindo $a = 10$, $b = 10000$ e $f_m = 20000Hz$.

A Fig. 6.3 representa, graficamente, os valores esperados de primeira (figura acima) e segunda (figura abaixo) ordem da variável $\Delta T(n)$. Foi verificado que para um número de realizações N superior a 400, o valor de $E\{\Delta T(n)\}$ converge para $\underline{\Delta T} = 1/20000s$, conforme frequência de amostragem, $f_m = 20000Hz$, estabelecida no código. Assim, para este modelo de densidade espectral de potência de $Y(t)$, utilizaremos as soluções da equação de Itô encontradas a partir de $N = 400$.

Uma vez definidos os parâmetros do sistema, foram simulados os possíveis períodos fundamentais de cada pulso glotal gerado. Uma forma de verificação da presença de *jitter* consiste na construção da função densidade de probabilidade (fdp) da variável aleatória definida por $F_{0g} = 1/T_g$, que será chamada de frequência fundamental. A Figura 6.4 mostra a fdp de F_{0g} para diferentes valores de a e b .

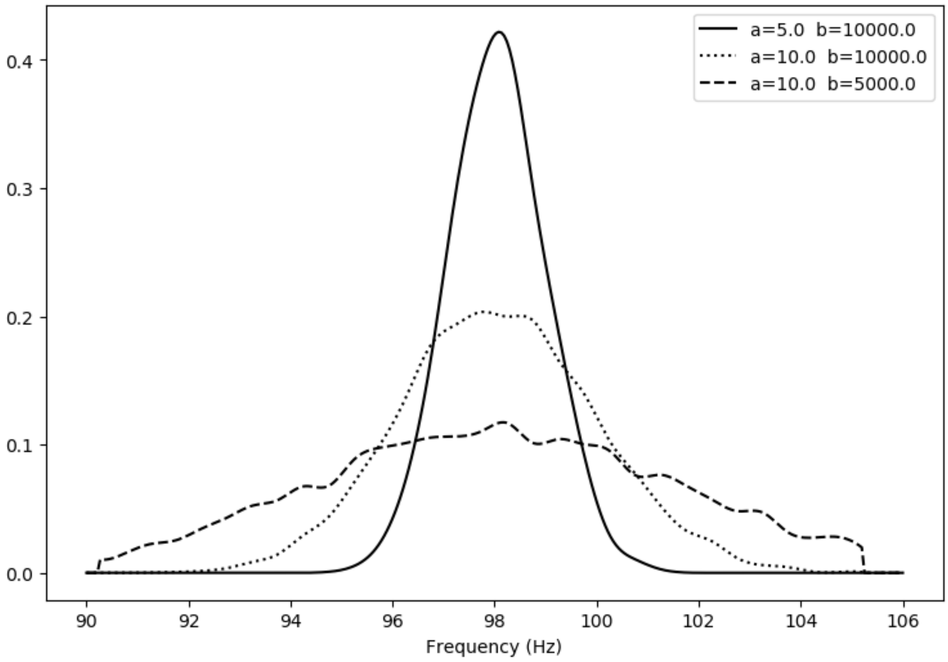


Figura 6.4: Fdp de F_{0g} para $a = 5.0$ e $b = 10000.0$; $a = 10.0$ e $b = 10000.0$; e $a = 10.0$ e $b = 5000.0$.

Através da Fig. 6.4, percebe-se que o aumento do parâmetro a e/ou a redução do parâmetro b acarreta na dispersão dos comprimentos dos períodos em torno do seu valor médio, ocasionando aumento no *jitter* produzido. Nas Fig. 6.5, 6.6, 6.7, 6.8 e 6.9, é possível observar simulações dos trens de pulsos de Rosenberg gerados para alguns casos, nos quais verifica-se o caso determinístico em $a = 0.0$ (Fig. 6.5).

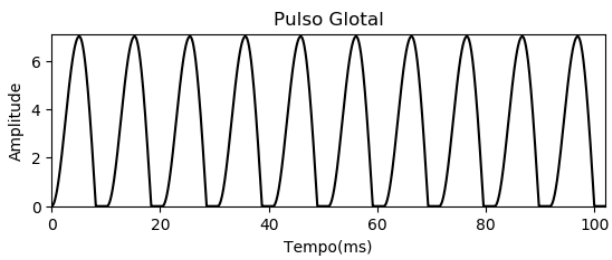


Figura 6.5: Trem de pulsos de Rosenberg - Caso 1: a=0.0 e b=10000.0 com $Jit_{loc} = 0.0\%$.

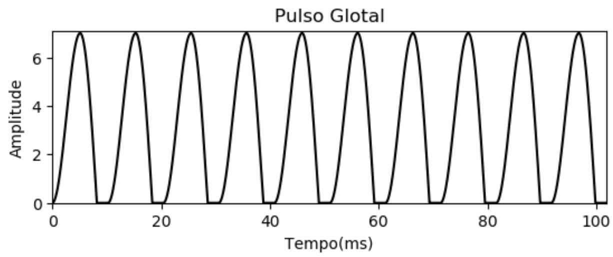


Figura 6.6: Trem de pulsos de Rosenberg - Caso 2: a=5.0 e b=10000.0 com $Jit_{loc} = 0.534\%$.

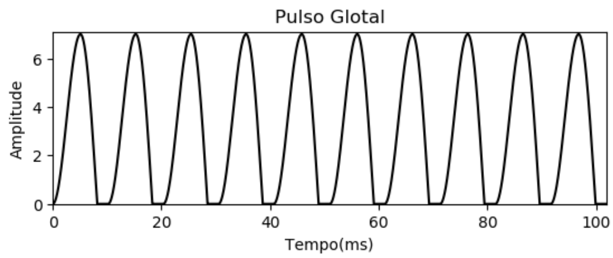


Figura 6.7: Trem de pulsos de Rosenberg - Caso 3: a=10.0 e b=10000.0 com $Jit_{loc} = 0.993\%$.

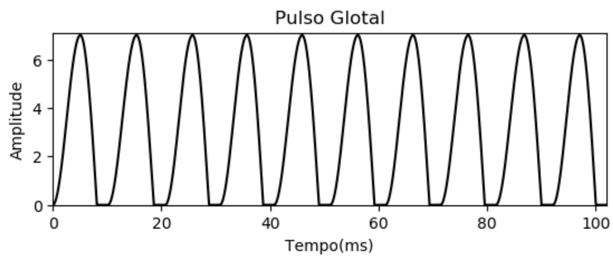


Figura 6.8: Trem de pulsos de Rosenberg - Caso 4: a=10.0 e b=5000.0 com $Jit_{loc} = 2.192\%$.

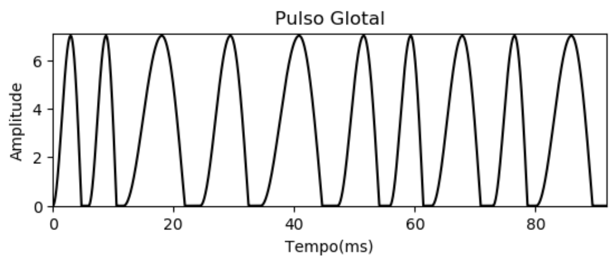


Figura 6.9: Trem de pulsos de Rosenberg - Caso 5: a=14.1 e b=100.0 com $Jit_{loc} = 43.26\%$.

Cada simulação nos dá indícios de que o *jitter* produzido está sendo controlado. Há casos com e sem indícios de patologias nas simulações executadas. Apesar da Fig. 6.9 indicar um valor elevado de flutuações nos períodos glotais, ele não representa um caso passível de ocorrência em humanos.

Os arquivos de áudio das vogais, sintetizados com diferentes níveis de *jitter*, podem ser ouvidos em: https://www.dropbox.com/sh/3c6n9sssnlvscf6/AADICuVloBFB8-SlvtMNjb9ka?dl=0 .

6.2 Caso II: Modelo de Rosenberg e densidade espectral a 3 parâmetros

Para este modelo de Rosenberg, foi aplicado o processo estocástico com densidade espectral de $Y(t)$ a três parâmetros, o que resultou na fdp indicada na Fig. 6.10. Percebe-se que a redução da variável a ocasiona uma redução na dispersão dos valores dos períodos fundamentais dos pulsos gerados em relação ao seu valor médio, e consequente diminuição do efeito *jitter*. O mesmo acontece com o aumento da variável c_b, apesar deste aumento implicar no deslocamento gráfico da densidade espectral de potência associada ao $Y(t)$ para a direita, conforme foi mostrado no capítulo anterior.

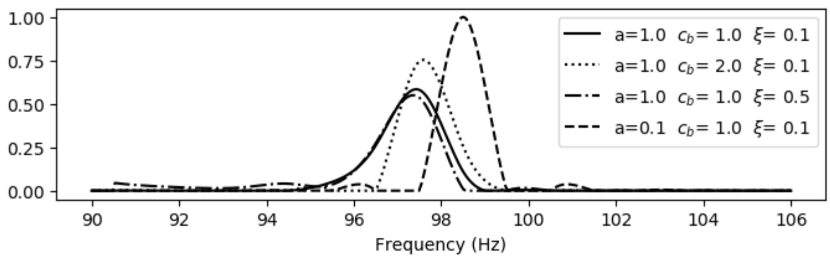


Figura 6.10: Fdp de F_{0g} para $a = 1.0$, $c_b = 1.0$ e $\xi = 0.1$; $a = 1.0$, $c_b = 2.0$ e $\xi = 0.1$; $a = 1.0$, $c_b = 1.0$ e $\xi = 0.5$; e $a = 0.1$, $c_b = 1.0$ e $\xi = 0.1$.

Nas Fig. 6.11, 6.12 e 6.13, aparecem 3 diferentes casos para demonstrar o controle do *jitter* através das variáveis envolvidas no processo estocástico estabelecido. É notório que o aumento de a ocasiona aumentos consideráveis do *jitter* no sistema, verificando-se o caso determinístico em $a = 0.0$ (Fig. 6.11).

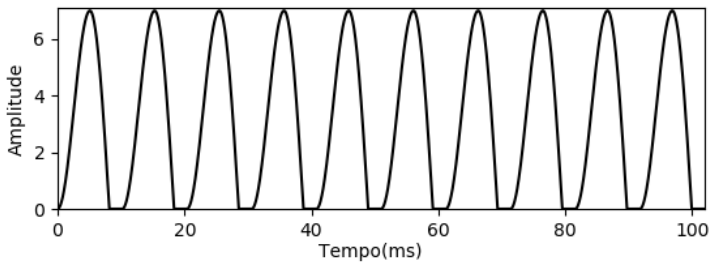


Figura 6.11: Trem de pulsos de Rosenberg - Caso 1: $a = 0.0$, $c_b = 1.0$ e $\xi = 0.1$ com $Jit_{loc} = 0.0\%$.

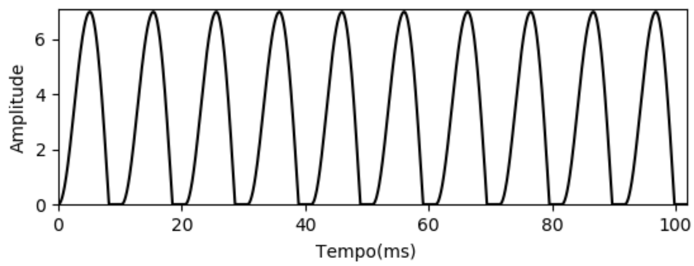


Figura 6.12: Trem de pulsos de Rosenberg - Caso 2: $a = 0.2$, $c_b = 1.0$ e $\xi = 0.1$ com $Jit_{loc} = 1.802\%$.

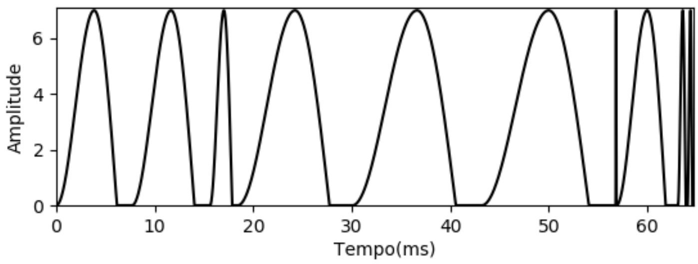


Figura 6.13: Trem de pulsos de Rosenberg - Caso 3: $a = 10.0$, $c_b = 1.0$ e $\xi = 0.1$ com $Jit_{loc} = 70.39\%$.

Os arquivos de áudio das vogais, sintetizados com diferentes níveis de *jitter*, podem ser ouvidos em: https://www.dropbox.com/sh/ggqminjjr51o2cj/AAA7PSaO9ndtIEPhufzTEBFga?dl=0 .

6.3 Caso III: Modelo de LF e densidade espectral a 2 parâmetros

Neste modelo, os períodos fundamentais dos pulsos glotais simulados comportam-se de forma similar ao primeiro caso apresentado (Modelo de Rosenberg e densidade espectral a 2 parâmetros). Assim, as variáveis a e b comportam-se do mesmo jeito, no que tange ao controle do fenômeno *jitter* gerado na voz sintetizada. A fdp para o caso em questão está representada na Fig. 6.14.

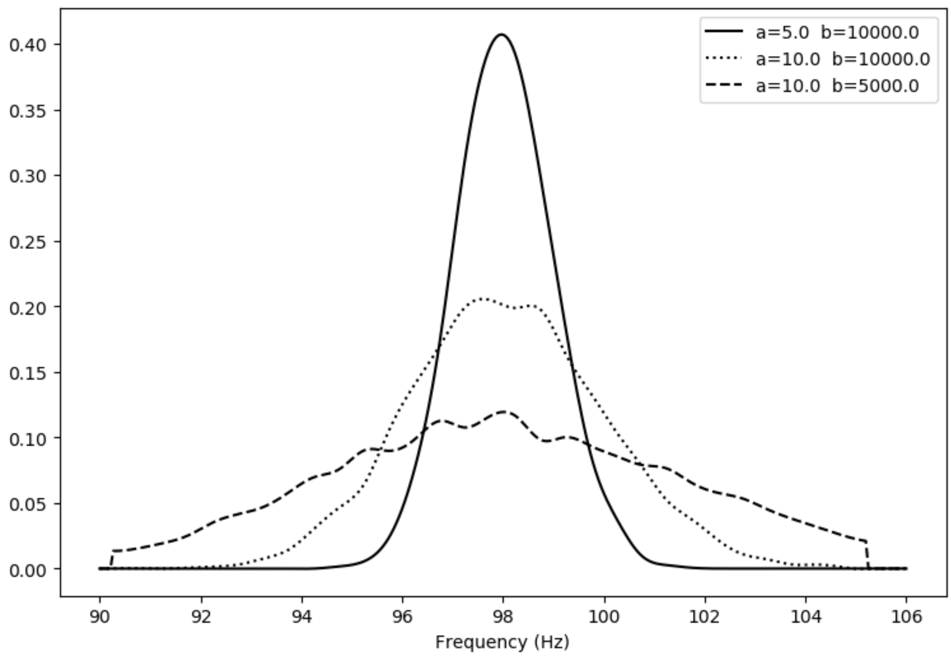


Figura 6.14: Fdp de F_{0g} para $a = 5.0$ e $b = 10000.0$; $a = 10.0$ e $b = 10000.0$; e $a = 10.0$ e $b = 5000.0$.

Nas Fig. 6.15, 6.16, 6.17, 6.18 e 6.19, estão representados graficamente os trens de pulsos com o modelo de LF para diferentes níveis de *jitter*, através da manipulação das variáveis a e b que controlam o fenômeno no sistema.

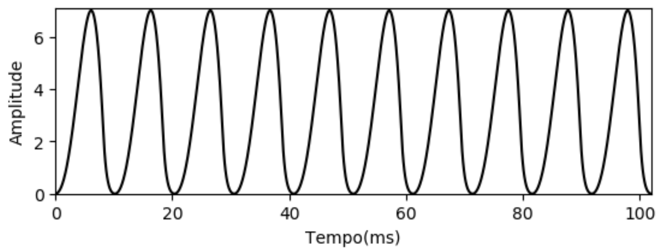


Figura 6.15: Trem de pulsos de LF - Caso 1: a=0.0 e b=10000.0 com $Jit_{loc} = 0.0\%$.

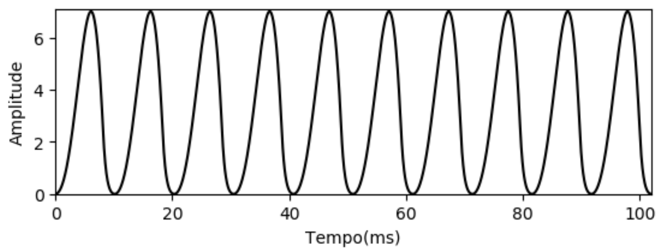


Figura 6.16: Trem de pulsos de LF - Caso 2: a=5.0 e b=10000.0 com $Jit_{loc} = 0.552\%$.

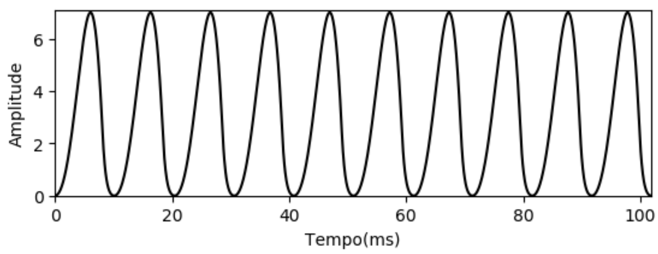


Figura 6.17: Trem de pulsos de LF - Caso 3: a=10.0 e b=10000.0 com $Jit_{loc} = 1.096\%$.

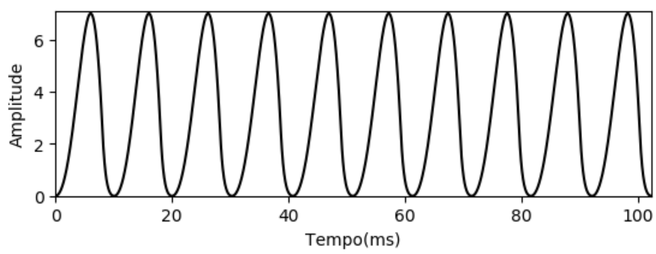


Figura 6.18: Trem de pulsos de LF - Caso 4: a=10.0 e b=5000.0 com $Jit_{loc} = 2.236\%$.

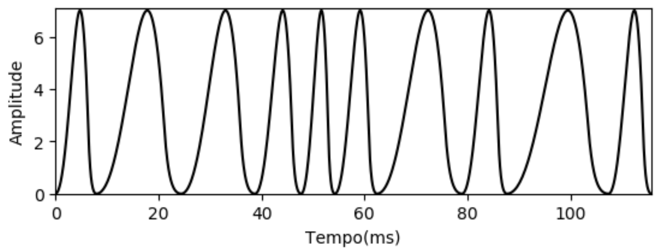


Figura 6.19: Trem de pulsos de LF - Caso 5: a=14.1 e b=100.0 com $Jit_{loc} = 53.66\%$.

Os arquivos de áudio das vogais, sintetizados com diferentes níveis de *jitter*, podem ser ouvidos em: https://www.dropbox.com/sh/la64ukd4u005wp2/AACxOqVZyBdY46AcDn-Gl9fNa?dl=0 .

6.4 Caso IV: Modelo de LF e densidade espectral a 3 parâmetros

Para este quarto caso, a Fig. 6.20 traz as funções densidade de probabilidade referentes a 4 casos representativos, a fim de mostrar o comportamento da frequência dos pulsos com a variação dos parâmetros, o qual apresentou comportamento similar ao segundo caso (Modelo de Rosenberg e densidade espectral a 3 parâmetros).

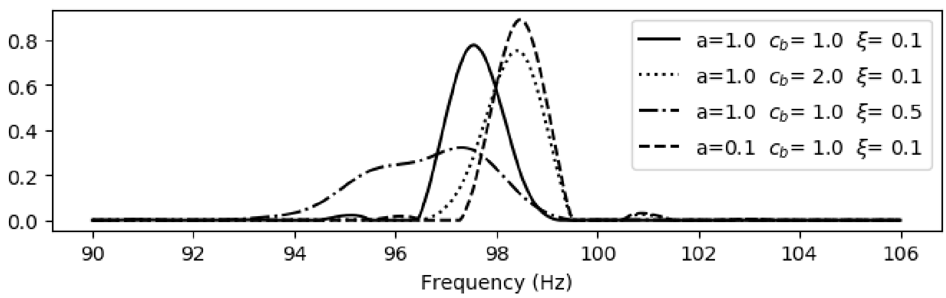


Figura 6.20: Fdp de F_{0g} para $a = 1.0$, $c_b = 1.0$ e $\xi = 0.1$; $a = 1.0$, $c_b = 2.0$ e $\xi = 0.1$; $a = 1.0$, $c_b = 1.0$ e $\xi = 0.5$; e $a = 0.1$, $c_b = 1.0$ e $\xi = 0.1$.

Nas Fig. 6.21, 6.22 e 6.23, os trens de pulsos gerados atendem às características do sistema, uma vez que a densidade espectral de potência de $Y(t)$ e as variáveis associadas a ela (a, c_b e ξ) conseguem, com eficiência, controlar o nível de *jitter* gerado na voz sintetizada.

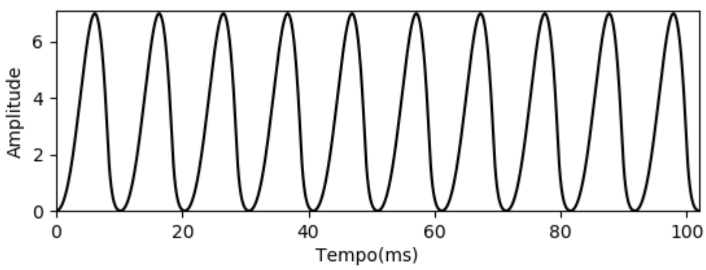


Figura 6.21: Trem de pulsos de LF - Caso 1: $a = 0.0$, $c_b = 1.0$ e $\xi = 0.1$ com $Jit_{loc} = 0.0\%$.

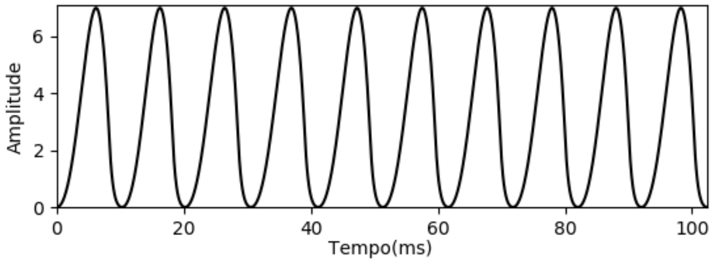


Figura 6.22: Trem de pulsos de LF - Caso 2: $a = 0.2$, $c_b = 1.0$ e $\xi = 0.1$ com $Jit_{loc} = 1.933\%$.

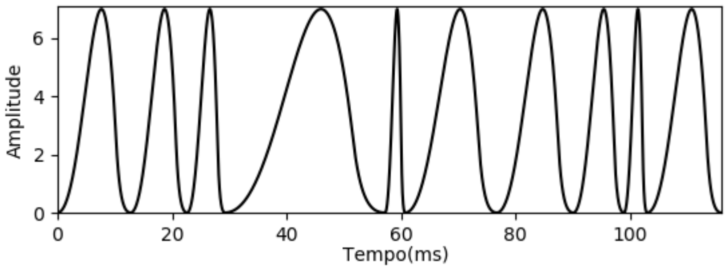


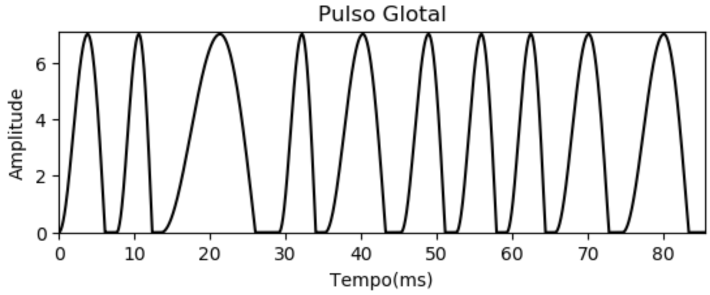
Figura 6.23: Trem de pulsos de LF - Caso 3: $a = 10.0$, $c_b = 1.0$ e $\xi = 0.1$ com $Jit_{loc} = 80.26\%$.

Os arquivos de áudio das vogais, sintetizados com diferentes níveis de *jitter*, podem ser ouvidos em: https://www.dropbox.com/sh/gv561zsh9855v1e/AACA0Gvrd1EMLqEvRIoyvRRka?dl=0 .

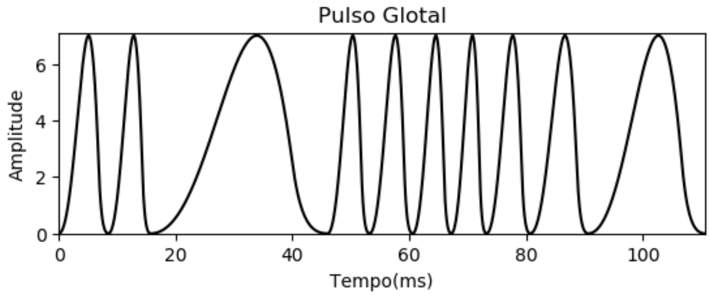
6.5 Comparação dos resultados

Nesta seção, será analisado o comportamento do sistema com a variação dos parâmetros que definem os modelos estocásticos implementados. Apesar de nos casos apresentados anteriormente já termos verificado o controle do efeito *jitter* através da manipulação das variáveis envolvidas, é importante que sejam mapeados os efeitos de cada alteração efetuada.

Na Fig. 6.24, temos a representação gráfica no tempo dos pulsos glotais com *jitter* para $f_0 = 98Hz$, levando em conta os modelos de pulsos de Rosenberg e LF. Como o período de amostragem dos pulsos $\Delta T(t)$, no caso com *jitter*, varia continuamente, isso proporciona mudanças do período de cada ciclo no trem de pulsos. Foi escolhido este caso, que possui um *jitter* local bem acima dos valores de *jitter* encontrados em humanos, apenas a fim de exemplificar variações críticas dos períodos dos pulsos ciclo a ciclo. Contudo, para o objetivo deste trabalho, serão escolhidos parâmetros que se adequem ao sistema humano, gerando um fenômeno *jitter* que se pareça ao apresentado por indivíduos saudáveis e também com indícios de patologias. O comprimento de cada pulso sofre uma expansão ou contração, devido à inserção do processo estocástico no sistema, tendo em $a = 0$ o caso determinístico, no qual não há *jitter*.



(a) Trem de Pulsos de Rosenberg com $Jit_{loc} = 50.36\%$.



(b) Trem de Pulsos de LF com $Jit_{loc} = 48.83\%$.

Figura 6.24: Trens de pulsos com jitter ($a = 14.1$ e $b = 100.0$).

Em relação ao modelo de densidade espectral a dois parâmetros, foram simuladas as funções densidades de probabilidades (fdp) da variável aleatória F_{0g}, inverso do período fundamental de cada ciclo glotal, mostrando como se comporta $\Delta t(t)$ como processo estocástico associado à densidade espectral de potência de $Y(t)$ com dois parâmetros, tomando diferentes valores de a e b. Por meio de simulações de Monte Carlo, foi possível definir as fdp da variável F_{0g}, representadas a seguir nas Fig. 6.25 e Fig. 6.26 .

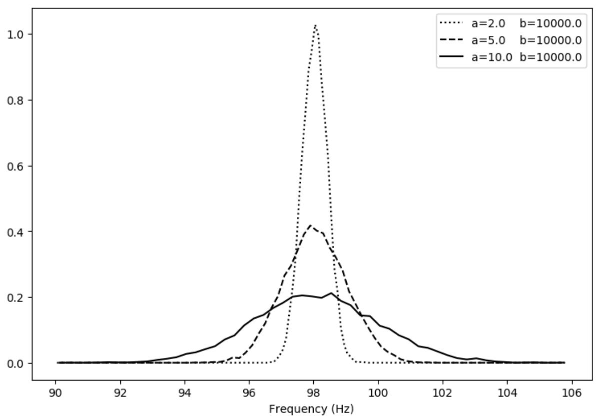


Figura 6.25: Fdp de F_{0g} para $b = 10000$, nos casos de $a = 2$, $a = 5$ e $a = 10$.

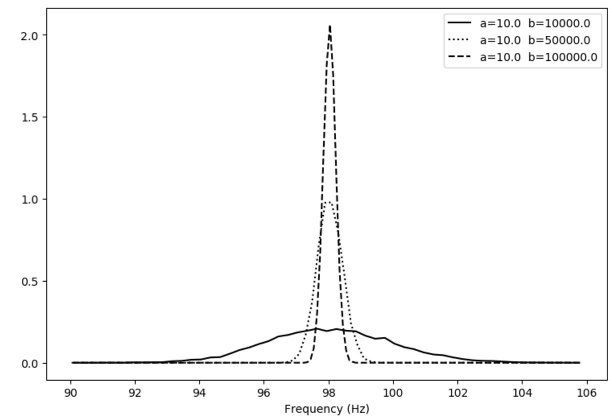


Figura 6.26: Fdp de F_{0g} para $a = 10$, nos casos de $b = 10000$, $b = 50000$ e $b = 100000$.

Foi mensurado o *jitter* local através da razão entre a média das diferenças absolutas entre períodos consecutivos e o período médio dos pulsos, dado pela Eq. 4.2. Como representado na Tab. 6.2, constata-se que valores maiores de b levam ao decréscimo do *jitter* local.

Casos	*jitter* Local
a=10 e b=100	46.85%
a=10 e b=1000	10.04%
a=10 e b=10000	1.016%
a=10 e b=100000	0.096%
a=10 e b=1000000	0.011%

Tabela 6.2: Valores de *jitter* locais simulados para $a = 10$ constante.

Como representado na Tab. 6.3, temos que valores maiores de a levam ao aumento do *jitter* local. É preciso encontrar valores ótimos para análise de sensibilidade e simulação de vozes próximas das emitidas por humanos, admitindo assim valores de *jitter* passíveis de ocorrência.

Casos	*jitter* Local
a=0 e b=10000	0.0%
a=2 e b=10000	0.222%
a=5 e b=10000	0.600%
a=10 e b=10000	1.016%
a=100 e b=10000	8.483%

Tabela 6.3: Valores de *jitter* locais simulados para $b = 10000$ constante.

Constatam-se então, a partir das Fig. 6.25 e 6.26 e das Tab. 6.2 e 6.3, que o fenômeno *jitter* está sendo controlado no sistema e que o aumento de a e/ou redução de b têm, juntos, gerado acréscimo nas flutuações do período fundamental em torno do seu valor médio.

Já em relação ao modelo de densidade espectral a três parâmetros, é possível identificar uma característica importante no que tange à variável a, uma vez que esta, ao passo que aumenta, tem implicado no aumento do efeito *jitter* da voz simulada, o que pode ser percebido na Tab. 6.4 .

Casos			
a	c_b	ξ	*jitter* Local
0.0	1.0	0.1	0.0%
1.0	1.0	0.1	0.109%
3.0	1.0	0.1	1.022%
5.0	1.0	0.1	3.126%
10.0	1.0	0.1	9.722%
50.0	1.0	0.1	87.33%

Tabela 6.4: Valores de *jitter* locais simulados para diferentes valores de a.

No que se refere à variável c_b, percebe-se através da Fig. 5.2 que ela tem relação direta com a frequência em que a função densidade espectral de potência apresenta valor máximo. Outrossim, percebe-se um decréscimo do efeito *jitter* com o aumento da referida variável, como pode ser observado na Tab. 6.5.

Casos			
a	c_b	ξ	*jitter* Local
3.0	1.0	0.1	1.022%
3.0	1.5	0.1	0.169%
3.0	2.0	0.1	0.067%
3.0	5.0	0.1	0.001%

Tabela 6.5: Valores de *jitter* locais simulados para diferentes valores de c_b.

A Tab. 6.6, a seguir apresentada, foi utilizada para mostrar que o aumento de ξ ocasionou redução no efeito *jitter*. A Fig. 5.1 mostra que, dados os outros parâmetros fixados, o aumento de ξ está relacionado com o aumento da banda de passagem da densidade espectral de potência do processo estocástico implementado.

Casos			
a	c_b	ξ	*jitter* Local
3.0	1.0	0.1	1.022%
3.0	1.0	0.3	1.006%
3.0	1.0	0.5	0.952%

Tabela 6.6: Valores de *jitter* locais simulados para diferentes valores de ξ.

Constatam-se então, a partir das Tab. 6.4, 6.5 e 6.6, que o fenômeno *jitter* está sendo controlado no sistema e que o aumento de a ou a redução de c_b ou ξ têm, juntos, gerado acréscimo nas flutuações do período fundamental em torno do seu valor médio.

A fim de melhorar a qualidade dos sons sintetizados, foi aplicada uma técnica de modulação da amplitude do sinal de saída, o que forneceu aos áudios uma característica atenuadora no início e no final da pronúncia de cada vogal. Assim, foram modificadas as frequências fundamentais dos sinais glotais e gerados os sons das vogais com um efeito *jitter* baixo, sem indícios de patologia. Os arquivos de áudio das vogais, sintetizados com diferentes frequências fundamentais, podem ser ouvidos em: https://www.dropbox.com/sh/dcwynhpxcvptakj/AAAbsF0Ny274V3YahenvpwUYa?dl=0 .

Em resumo, constatam-se as seguintes informações no que tange à comparação dos modelos:

1. Modelo de pulso glotal de Rosenberg e densidade espectral a 2 parâmetros: Modelo estocástico que permite controle simples do formato de pulso, uma vez que considera fechamento abrupto das cordas vocais, e controle moderado do *jitter*, dado que modifica apenas a intensidade do *jitter* através de seus parâmetros a e b;

2. Modelo de pulso glotal de Rosenberg e densidade espectral a 3 parâmetros: Modelo estocástico que permite controle simples do formato de pulso, uma vez que considera fechamento abrupto das cordas vocais, e controle significativo do *jitter*, dado que também modifica a frequência fundamental e o colorido frequencial do sinal de voz gerado através de seus parâmetros a, c_b e ξ;

3. Modelo de pulso glotal de LF e densidade espectral a 2 parâmetros: Modelo estocástico que permite controle mais complexo do formato de pulso, uma vez que permite definir a fase de retorno como uma função exponencial decrescente, e controle moderado do *jitter*, dado que modifica apenas a intensidade do *jitter* através de seus parâmetros a e b; e

4. Modelo de pulso glotal de LF e densidade espectral a 3 parâmetros: Modelo estocástico que permite controle mais complexo do formato de pulso, uma vez que permite definir a fase de retorno como uma função exponencial decrescente, e controle significativo do *jitter*, dado que também modifica a frequência fundamental e o colorido frequencial do sinal de voz gerado através de seus parâmetros a, c_b e ξ.

Percebe-se, também, no que tange à comparação dos áudios da síntese das vogais com diferentes excitações glotais, que a fonte glotal do modelo de Rosenberg gera um som metalizado com maior energia acústica do que o modelo LF, que possui uma qualidade mais aflautada no som emitido.

Ademais, uma avaliação importante é que, através dos arquivos de áudio obtidos, constata-se a qualidade rouca e áspera associada aos sons, à medida que o fenômeno *jitter* era acrescido e superior a 1% nas vogais sintetizadas. Para variações muito pequenas do período fundamental na excitação glotal, percebe-se a emissão de sons vocálicos mais naturais e, por conseguinte, sem indícios de presença de patologias.

Vale ressaltar que para a geração de vogais mais próximas das que são emitidas pelo ser humano, é indispensável que o trem de pulsos seja o mais próximo possível do que é produzido na glote, além da dependência direta dos formantes e larguras de banda definidas no código.

Capítulo 7

Conclusões e sugestões para trabalhos futuros

7.1 Conclusões

A modelagem estocástica do sinal glotal, a partir do modelo LF e de Rosenberg, mostrou-se eficiente para gerar o *jitter* nos sinais de voz. Foi possível obter sinais com diferentes níveis de *jitter* e, principalmente, níveis que podem indicar patologias das cordas vocais. O *jitter* por si só, mesmo abaixo do nível de patologia, está presente em todas as vozes e ajuda a fornecer um caráter mais natural à voz sintetizada.

Dentro do mesmo modelo de *jitter*, é possível perceber que o som gerado pelo modelo de Rosenberg é mais metalizado, com maior energia acústica quando comparado ao de LF, o qual é mais suave, mais com estilo de flauta, com diferentes propriedades, o que permite inferir que a forma de onda do pulso glotal está interligada não somente à energia como também à qualidade do som gerado.

Conforme era esperado, identificou-se que na voz gerada com *jitter*, à medida que o Jit_{loc} atingia valor superior a 1% nas vogais sintetizadas, o som ganhou uma característica adicional que o tornava áspero e rouco, podendo ser sinal de alguma patologia.

Para o modelo de densidade a dois parâmetros, verificou-se que o aumento do valor do parâmetro a acarreta no aumento do nível de *jitter*, ao contrário da variável b. Uma vez escolhido o valor de b, o parâmetro a atua como um ajuste fino, estabelecendo valores de *jitter* de acordo com as pequenas variações desejadas.

A comparação entre as funções de densidade de probabilidade da frequência fundamental para os vários casos da densidade espectral a duas variáveis permite perceber que o *jitter* foi gerado, como inicialmente proposto.

Para as densidades espectrais a três variáveis, há um número maior de possibilidades em torno do período fundamental estabelecido, uma vez que os parâmetros do sistema determinam como as frequências fundamentais se comportam, possibilitando modificar o *colorido* das frequências no sinal. Percebe-se que a variação dos parâmetros a, c_b e ξ implica diretamente no controle do *jitter*, sendo que o Jit_{loc} possui maior sensibilidade com os aumentos e decréscimos de a. É notado que aumentos em c_b ($b = 2\pi c_b f_p$) proporcionam mudanças na frequência correspondente ao valor máximo da curva de $S_Y(\omega)$, deslocando-a pra direita. Dessa forma, conclui-se que com esse modelo estocástico é possível mudar a frequência de voz estabelecida no sistema. Já a variável ξ mostra-se correlacionada positivamente com a banda de passagem da densidade espectral de potência do processo estocástico implementado. O controle do colorido frequencial permitido por este modelo traz inúmeras vantagens ao sistema. Assim, faz-se necessário propor a solução de um problema inverso estocástico para identificar os parâmetros/processos estocásticos do modelo criado.

O método utilizado neste trabalho é inédito e possibilita controlar, além do nível de *jitter*, também a faixa de frequências a ser utilizada. Além disso, considera um processo estocástico diferente do gaussiano, que nesse caso não se aplicaria e seria fisicamente incorreto.

7.2 Sugestões para trabalhos futuros

Como sugestão, tem-se a criação de um problema estocástico inverso para identificação dos parâmetros estabelecidos no modelo criado neste trabalho e posterior tratamento estatístico das informações coletadas. Para isso, faz-se necessário que seja implementado um sistema reverso para extração do pulso glotal, a partir do som emitido pelo ser humano, e criada uma amostra satisfatória de indivíduos saudáveis e com presença de patologias no trato vocal, dos quais serão colhidos os sons vocálicos.

Referências Bibliográficas

[1] E. Cataldo e E. Barrientos, "Identificação de parâmetros em um modelo matemático para geração da voz humana", *Proceeding Series of the Brazilian Society of Applied and Computational Mathematics*, Vol. 6, n. 1, 2018.

[2] E. Cataldo and C. Soize, "Voice signals produced with *jitter* through a stochastic one-mass mechanical model", *Journal of voice*, Vol. 31, pp. 111.e9-111.e18, 2017.

[3] J. Schoentgen, "Stochastic models of *jitter*", *The Journal of the Acoustical Society of America*, Vol. 109, n. 4, pp. 1631-1650, 2001.

[4] D. Ruinskiy and Y. Lavner, "Stochastic models of pitch *jitter* and amplitude shimmer for voice modification", in *IEEE 25th Convention of Electrical and Electronics Engineers in Israel*, pp. 489-493, 2008.

[5] E. Cataldo and C. Soize, "Stochastic mechanical model of vocal folds for producing *jitter* and for identifying pathologies through real voices", *Journal of Biomechanics*, Vol. 74, pp. 126-133, 2018.

[6] D. G. Silva, L. C. Oliveira and M. Andrea, "*jitter* Estimation Algorithms for Detection of Pathological Voices", *EURASIP Journal on Advances in Signal Processing*, n. 567875, Lisboa, Portugal, 2009.

[7] J.P. Dworkin, R.J. Meleca, "Vocal Pathologies: Diagnosis, Treatment and Case Studies", 2nd edition, New York, 1997. 336 p.

[8] J. Kreiman, B.R. Gerratt, G.B. Kempster, A. Erman, and G.S. Berke, "Perceptual Evaluation of Voice Quality: Review, Tutorial, and a Framework for Future Research", *J. Speech Hear. Res.*, Vol. 36, pp. 21-40, 1993.

[9] W. R. Zemlin, "Anatomia e Fisiologia aplicada a fonoaudiologia", 4. ed., Porto Alegre: Artmed, 2000. 624 p.

[10] Aparelho fonador. Fonética e Fonologia: Sonoridade em Artes, Saúde e Tecnologia. [Online] Disponível em: http://fonologia.org/aparelho_fonador.php. [Acessado em 05 de fevereiro de 2020].

[11] Portal Educação. O aparelho fonador e os mecanismos da produção dos sons. [Online] Disponível em: https://www.portaleducacao.com.br/fonoaudiologia/artigos/23463/o-aparelho-fonador-e-os-mecanismos-de-producao-dos-sons. [Acessado em 05 de fevereiro de 2020].

[12] J. D. Vieira, "Laringe: Falsas Cordas Vocais e as Cordas Verdadeiras", *ACM - Arquivos Catarinenses de Medicinaa*, Vol. 32, pp. 1-2, Florianópolis, 2003.

[13] R. L. Gottardo, "Um modelo físico simples para a descrição do funcionamento das pregas vocais", Trabalho de Final de Curso, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.

[14] R. C. Meirelles, R. Bak e F. C. Cruz, "Presbifonia", *Revista HUPE-UERJ*, Vol. 11, n. 3, pp. 77-82, Rio de Janeiro, 2012.

[15] Infecções Respiratórias Agudas (IRA): Trato Respiratório Inferior. [Online] Disponível em: http://bioemfoco.com.br/noticia/infeccoes-respiratorias-agudas-ira-trato-respiratorio-inferior/. [Acessado em 07 de janeiro de 2020].

[16] F. M. Coelho, "Modelagem empírica do trato vocal e sua influência no timbre da flauta", Trabalho apresentado no IV Seminário de Música Ciência e Tecnologia: Fronteiras e Rupturas, Universidade de São Paulo, São Paulo, 2012.

[17] Y. Hori, "Fundamental frequency pertubation observed in sustained phonation", *J. Speech Hear. Res.*, Vol. 22, pp. 5-19, 1979.

[18] L. C. Cagliari, "Elementos de Fonética do português brasileiro", Tese de livre docência, Unicamp, Campinas, São Paulo, 1981.

[19] G. Fant, "Acoustic theory of speech production", 2nd edition, Mouton, The Hague, 1970. 328 p.

[20] L. R. Rabiner and R. W. Shafer, "Digital Processing of Speech Signals", Prentice-Hall, 1978. 962 p.

[21] G. Degottex, "Glottal Source and Vocal-Tract Separation - Estimation of glottal parameters, voice transformation and synthesis using a glottal model", Tese de doutorado, Université Paris, 2010.

[22] G. Fant, "Vocal-Source Analysis - A Progress Report", *STL-QPSR*, nos. 3-4, pp. 31-53, 1979.

[23] J. Liljencrants, G. Fant and Q. Lin, "A Four Parameter Model of Glottal Flow", *STL-QPSR*, n. 4, pp. 1-13, 1985.

[24] L. C. Klatt, "Analysis, Synthesis, and Perception of Voice Quality Variations among Female and Male Talkers", *Journal of the Acoustical Society of America*, n. 87, pp. 820-857, 1990.

[25] A. C. Guerra, "Estimação do sinal glotal para padrões acústicos de doenças da laringe", Tese de Doutorado, Universidade de São Paulo, Escola de Engenharia de São Carlos, São Paulo, 2005.

[26] N. H. Bernardoni, C. d'Alessandro e B. Doval, "Glottal flow models: waveforms, spectra and physical measurements", *Forum Acusticum*, Sevilla, Spain, p. 1, 2002.

[27] B. Doval, C. d'Alessandro, N. Henrich, "The spectrum of glottal flow models", *Acta acustica united with acustica*, Vol. 92, n. 6, pp. 1026-1046, 2006.

[28] M. Behlau, G. Madazio, D. Feijó e P. Pontes, "Avaliação de Voz", *In: BEHLAU*, Mara. Voz o livro do especialista, Vol. 1, Rio de Janeiro: Revinter. Cap. 3, pp. 130-164, Rio de Janeiro, 2001.

[29] M. Oliveira e V. Pacheco, "Características acústicas da vogal /i/ produzida por sujeitos com Síndrome de Down", *Revista Veredas*, Vol. 16, n. 2, pp. 108-109, 2012.

[30] L. E. B. Sandoval e E. Cataldo, "Comparação de modelos de sinal glotal na síntese de vogais, nos casos de vogal sustentada e de voz cantada, considerando sons na língua espanhola", Dissertação de Mestrado, Universidade Federal Fluminense, Niterói, Rio de Janeiro, 2018.

[31] R. D. Kent e C. Read, "Análise Acústica da fala", Tradução de Alexsandro Rodrigues Meireles, 1a. edição, São Paulo, Cortez, 2015. 503 p.

[32] Software Praat. Universidade de Amsterdam. [Online] Disponível em: http://www.fon.hum.uva.nl/praat/. [Acessado em 24 de agosto de 2019].

[33] M. K. Christmann, A. R. Brancalioni, C. R. Freitas, D. Z. Vargas, M. Keske-Soares, C. L. Mezzomo e H. B. Motai, "Uso do programa MDVP em diferentes contextos: revisão de literatura", *Revista CEFAC*, Vol. 17, n. 4, pp. 1341-1349, 2015.

[34] D. G. Childers, "Speech Processing and Synthesis Toolboxes", Wiley, New York, pp. 390–393, 2000.

[35] N. Henrich, "Etude de la source glottique en voix parlée et chantée: modélisation et estimation, mesures acoustiques et électroglottographiques, perception", Apresentada como tese de doutorado, Université Pierre et Marie Curie-Paris VI (UPMC), 2001.

[36] C. Soize, "The Fokker-Planck Equation for Stochastic Dynamical Systems and its Explicit Steady State Solutions", *World Scientific*, Singapore, 1994.

[37] M. N. Vieira, "Uma introdução à acústica da voz cantada", *AcMUS - I Seminário Música, Ciência e Tecnologia*, Vol.1, pp. 70-79, São Paulo, 2004.

Apêndice A

Pulso unitário determinístico de Rosenberg

```
import matplotlib.pyplot as plt
import numpy as np
from numpy import sin,cos,exp
from scipy.optimize import fsolve
from math import pi
#Pulso Unitario Deterministico
def ros_uno_uni(Av,FO,fm,alpha1,alpha2):
  TO=1/float(FO)
  Tm=1/float(fm)
  TP =(alpha1/float(100))*TO
  TN =(alpha2/float(100))*TO
  t1 = np.arange(0,TP,Tm)
  t2 = np.arange(TP,TP+TN,Tm)
  t12= np.concatenate((t1,t2))
  t3 = np.arange(TP+TN,TO,Tm)
  t=np.concatenate((t1,t2,t3))
  x1 = Av*0.5*(1-np.cos(np.pi*t1/(TP)))
  x2 = Av*(np.cos(np.pi*(t2-TP)/(2*TN)))
  x3 = 0*t3
  x=np.concatenate((x1,x2,x3))
  y1 = Av*0.5*np.pi*(np.sin(np.pi*t1/(TP)))/TP
  y2 = -Av*np.pi*(np.sin(np.pi*(t2-TP)/(2*TN)))/(2*TN)
  y12=np.concatenate((y1,y2))
  y3 = 0*t3
  y=np.concatenate((y1,y2,y3))
  imp=np.zeros(int(TO/Tm)+2)
  imp[0]=1
  z = np.convolve(x,imp)
  return x,t,TO
```

Apêndice B

Pulso unitário estocástico de Rosenberg e densidade espectral a 2 parâmetros

```
# -*- coding: utf-8 -*-
import matplotlib.pyplot as plt
import numpy as np
from numpy import sin,cos,exp
from scipy.optimize import fsolve
from math import pi
#Pulso Unitario com Processo Estocastico
def ros_uno_uni_normal(Av,FO,fm,alpha1,alpha2,abar,bbar):
  Tm=1/float(fm) #intervalo de amostragem do pulso
  TO=1/float(FO) #periodo fundamental do ciclo
  TP =(alpha1/float(100))*TO
  TN =(alpha2/float(100))*TO
  xbar=(1.-abar**2/(2*bbar))**0.5
  kbar=10000
  beg=3000
  xx=([])
  xini=0
  ko=1/40000.
  for ii in range(kbar):
    xfin=([(((-float(bbar)*Tm/2.+1.)*xini+abar*np.random.randn()*Tm**0.5)/(1+float(bbar)*
        Tm/2.)])
    xx=np.concatenate((xx,xfin))
    xini=float(xfin[0])
  dT=ko+(Tm-ko)*(xbar+xx[beg:])**2
  TOfin=sum(dT[0:int(TO/Tm)])
  TOvetor=TOfin*np.ones(int(TO/Tm)+1)
  t=([0])
  for iii in range(len(dT)):
```

```
    t=np.concatenate((t,([t[-1]+dT[iii]])))
  t=t[0:int(TO/Tm)+1]
  x=([])
  k=0
  n=0
  TP =(alpha1/float(100))*TOfin
  TN =(alpha2/float(100))*TOfin
  for i in range(len(t)):
    if t[i]<=TP:
      k+=1
    if t[i]<=TP+TN:
      n+=1
  #criacao do pulso de Ros com deltat(t) processo estocástico
  for i in range(k):
    x1 = [(Av*0.5*(1-np.cos(np.pi*t[i]/(TP))))]
    x=np.concatenate((x,x1))
  for i in range(k,n):
    x2 = [(Av*(np.cos(np.pi*(t[i]-TP)/(2*TN))))]
    x=np.concatenate((x,x2))
  for i in range(n,len(t)):
    x3 = [(0)]
    x=np.concatenate((x,x3))

  y = np.diff(x)/Tm #derivada do pulso no tempo
  imp=np.zeros(int(TOfin/Tm)+2) #função impulso delta(t)
  imp[0]=1
  z = np.convolve(x,imp)

  return x,t,TOfin
```

Apêndice C

Trem de pulsos estocásticos de Rosenberg e densidade espectral a 2 parâmetros - Cálculo das medidas de dispersão

```
import os
import matplotlib.pyplot as plt
import numpy as np
from numpy import cos,exp, log10,pi
from scipy.signal import lfilter, freqz
import soundfile as sf
from Pulso_Unitario_Ros import ros_uno_uni_normal
from Pulso_Unitario_Ros import ros_uno_uni
from random import random

a=14.1
b=100.0
os.makedirs('a='+str(a)[0:6]+'_b='+str(b))
os.chdir('C:/Users/diego/OneDrive/Documentos/Mestrado_UFF\Codigos\duasvariaveis\
    Codigos_2_duasvariaveis/a='+str(a)[0:6]+'_b='+str(b))
fm=20000 #frequencia de amostragem
Tm=1/float(fm) #periodo de amostragem
Av=7 #amplitude de vozeamento
Tfs=3 #tempo total de sustentacao da vogal
FO=98 #frequencia funcamental
TO=1/float(FO) #periodo fundamental
alpha1=50.
alpha2=30.
x1=([])
g1=([])
```

```
jit =([])
todas =([])
t =([])
tz=0
#cria e concatena os pulsos
for i in range(int(Tfs/TO)+10):
  x,ty,TOfin=ros_uno_uni_normal(Av,FO,fm,alpha1,alpha2,a,b) #cria um pulso com processo
      estocastico
  x1=np.concatenate((x1,x))#concatena os pulsos
  tt=ty+tz
  t=np.concatenate((t,tt))
  tz=t[-1]
  jit=np.concatenate((jit,([TOfin])))
  x=x1
for i in range(int(Tfs/TO)+1):
  g,t1,TOo=ros_uno_uni(Av,FO,fm,alpha1,alpha2) #cria um pulso Ros determinístico
  g1=np.concatenate((g1,g)) #concatena os pulsos
g=g1
t1=np.arange(0,len(g)*Tm,Tm)
j=len(jit)
jitabs=0
for i in range(j-1):
  jitabs+=abs((jit[i]-jit[i+1]))
jitabs=jitabs/(j-1)
jitrel=jitabs/(sum(jit)/j)
jitrel=jitrel*100
print jitrel

#novos parâmetros
soma1=0
for i in range(1,j-1):
  soma1+=abs((2*jit[i]/3.-jit[i-1]/3.-jit[i+1]/3.))
soma1=soma1/(j-2)
rap=soma1/(sum(jit)/j)
rap=rap*100
soma2=0
for i in range(2,j-2):
  soma2+=abs((4*jit[i]/5.-jit[i-1]/5.-jit[i-2]/5.-jit[i+1]/5.-jit[i+2]/5.))
soma2=soma2/(j-4)
ppq5=soma2/(sum(jit)/j)
ppq5=ppq5*100
soma3=0
for i in range(1,j-1):
  soma1+=abs((jit[i+1]+jit[i-1]-2*jit[i]))
soma3=soma3/(j-2)
ddp=soma3/(sum(jit)/j)
ddp=ddp*100

arquivo = open('parametros','w')
```

```
arquivo.write('jit = '+str(jitrel)+'% \n')
arquivo.write('rap = '+str(rap)+'% \n')
arquivo.write('ppq5 = '+str(ppq5)+'% \n')
arquivo.write('ddp = '+str(ddp)+'% \n')
arquivo.close()

#criacao do delta de dirac
imp = np.zeros(len(t))
imp[0] = 1
#convolucao entre o delta de dirac e o Pulso Glotal de Ros deterministico
z2 = np.convolve(g,imp)
#convolucao entre o delta de dirac e o Pulso Glotal de Ros com Delta t(t) processo
    estocastico
z1 = np.convolve(x,imp)

#Trato Vocal
for H in range(1,6):
  L=['A','E','I','O','U']
  #formantes da categoria de cantor baixo
  formants=[[900,1300,2000,2200,2500],
    [450,1700,2000,2200,2310],
    [300,1900,2100,2200,2490],
    [500,800,2150,2200,2490],
    [360,700,2170,2200,2330]]
  bandwidths=[41,52,70,32,100]
  fmts=formants[H-1]
  if FO>fmts[0]:
    fmts[0]=FO+10
  impulse = np.zeros(500)
  impulse[0] = 1
  xin=impulse
  for resonance in range(5):
    f=fmts[resonance]
    bw=bandwidths[resonance]
    num=np.array([1-2*exp(-bw*2*pi/fm)*cos(2*pi*f/fm)+exp(-4*bw*pi/fm)])
    den=np.array([1,-2*exp(-bw*2*pi/fm)*cos(2*pi*f/fm),exp(-4*bw*pi/fm)])
    yout=lfilter(num,den,xin,axis=0)
    xin=yout
  #Radiacao dos Labios (Aproximacao)
  r=0.95
  B1=np.array([1,-r])
  A1=np.array([1])
  yout=lfilter(B1,A1,np.squeeze(xin))
  #convolucao entre os pulso glotais e os filtros do trato vocal
  zout=np.convolve(yout,z1)
  ym=zout[0:len(x)]
  #salvar em áudio cada vogal sintetizada pelo modelo fonte-filtro
  sf.write('Ros_'+str(L[H-1])+'_a='+str(a)+'_b='+str(b)+'_jit'+str(jitrel)[0:5]+'%.wav',
      ym, fm)
```

```
  #concatena um trecho de cada vogal sustentada para gerar um áudio com todas vogais
  ym=ym[0:len(ym)/5]
  ym=np.concatenate((ym,np.zeros(len(ym))))
  todas=np.concatenate((todas,ym))

sf.write('Todas_Ros_a='+str(a)+'_b='+str(b)+'_jit'+str(jitrel)[0:5]+'%.wav', todas, fm) #
    salva o áudio de todas as vogais em um arquivo
```

Apêndice D

Pulso unitário estocástico de Rosenberg e densidade espectral a 3 parâmetros

```
# -*- coding: utf-8 -*-
import matplotlib.pyplot as plt
import numpy as np
from numpy import sin,cos,exp
from scipy.optimize import fsolve
from math import pi
from numpy.linalg import inv

#Pulso Unitario com Processo Estocastico
def ros_uno_uni_normal(Av,FO,fm,alpha1,alpha2,abar,bbar,eta):
    alpha=np.array([[0.,1.],[-float(bbar)**2,-2*eta*float(bbar)]])
    beta=np.array([[0],[abar]])
    Tm=1/float(fm) #intervalo de amostragem do pulso
    TO=1/float(FO) #periodo fundamental do ciclo
    xbar=(1.-abar**2/(4*eta*bbar**3))**0.5
    TP =(alpha1/float(100))*TO
    TN =(alpha2/float(100))*TO
    kbar=10000
    beg=3000
    xx=([])
    xini=np.array([[0],[0]])
    ko=1/40000.
    for ii in range(kbar):
        xfin=(-(alpha)*Tm/2.+np.ones((2,2))).dot(xini)
        xfin=(xfin+beta*np.random.randn()*Tm**0.5)
        xfin=inv((np.ones((2,2))+alpha*Tm/2.)).dot(xfin)
        kfin=([xfin[0][0]])
        xx=np.concatenate((xx,kfin))
```

```
        xini=xfin
    dT=ko+(Tm-ko)*(xbar+xx[beg:])**2
    TOfin=sum(dT[0:int(TO/Tm)])
    TOvetor=TOfin*np.ones(int(TO/Tm)+1)
    t=([0])
    for iii in range(len(dT)):
        t=np.concatenate((t,([t[-1]+dT[iii]])))
    t=t[0:int(TO/Tm)+1]
    x=([])
    k=0
    n=0
    TP =(alpha1/float(100))*TOfin
    TN =(alpha2/float(100))*TOfin
    for i in range(len(t)):
        if t[i]<=TP:
            k+=1
        if t[i]<=TP+TN:
            n+=1

    #criacao do pulso de Ros com delta t(t) processo estocástico
    for i in range(k):
        x1 = [(Av*0.5*(1-np.cos(np.pi*t[i]/(TP))))]
        x=np.concatenate((x,x1))
    for i in range(k,n):
        x2 = [(Av*(np.cos(np.pi*(t[i]-TP)/(2*TN))))]
        x=np.concatenate((x,x2))
    for i in range(n,len(t)):
        x3 = [(0)]
        x=np.concatenate((x,x3))

    y = np.diff(x)/Tm #derivada do pulso no tempo
    imp=np.zeros(int(TOfin/Tm)+2) #função impulso delta(t)
    imp[0]=1
    z = np.convolve(x,imp) #convolucao entre o impulso unitario e o sinal do pulso que resulta no proprio sinal do pulso

    return x,t,TOfin
```

Apêndice E

Trem de pulsos estocásticos de Rosenberg e densidade espectral a 3 parâmetros - Cálculo das medidas de dispersão

```
import os
import matplotlib.pyplot as plt
import numpy as np
from numpy import cos,exp, log10,pi
from scipy.signal import lfilter, freqz
import soundfile as sf
from Pulso_Unitario_Ros import ros_uno_uni_normal
from Pulso_Unitario_Ros import ros_uno_uni
from random import random

a=0.15
c=1.0
eta=0.1
#print os.getcwd()
os.makedirs('a='+str(a)[0:6]+'_c='+str(c)+'_eta='+str(eta))
os.chdir('C:/Users/diego/OneDrive/Documentos/Mestrado_UFF/Codigos/tresvariaveis/Codigos_2_tresvariaveis/a='+str(a)[0:6]+'_c='+str(c)+'_eta='+str(eta))
fm=20000 #frequencia de amostragem
Tm=1/float(fm) #periodo de amostragem
Av=7 #amplitude de vozeamento
Tfs=3 #tempo total de sustentacao da vogal
FO=98 #frequencia funcamental
b=2*pi*FO*c
TO=1/float(FO) #periodo fundamental
alpha1=50.
```

```
alpha2=30.
x1=([]) #criacao do vetor de pulsos Ros com delta t(t) processo estocastico
g1=([]) #criacao do vetor de pulsos Ros Determinístico
jit=([])
todas=([])
t=([])
tz=0
#cria e concatena os pulsos
for i in range(int(Tfs/TO)+10):
  x,ty,TOfin=ros_uno_uni_normal(Av,FO,fm,alpha1,alpha2,a,b,eta) #cria um pulso com delta t(t) processo estocastico
  x1=np.concatenate((x1,x))#concatena os pulsos
  tt=ty+tz
  t=np.concatenate((t,tt))
  tz=t[-1]
  jit=np.concatenate((jit,([TOfin])))
x=x1
for i in range(int(Tfs/TO)+1):
  g,t1,TOo=ros_uno_uni(Av,FO,fm,alpha1,alpha2) #cria um pulso ros determinístico
  g1=np.concatenate((g1,g)) #concatena os pulsos
g=g1
t1=np.arange(0,len(g)*Tm,Tm)
j=len(jit)
jitabs=0
for i in range(j-1):
  jitabs+=abs((jit[i]-jit[i+1]))
jitabs=jitabs/(j-1)
jitrel=jitabs/(sum(jit)/j)
jitrel=jitrel*100
print jitrel
#novos parâmetros
soma1=0
for i in range(1,j-1):
  soma1+=abs((2*jit[i]/3.-jit[i-1]/3.-jit[i+1]/3.))
soma1=soma1/(j-2)
rap=soma1/(sum(jit)/j)
rap=rap*100
soma2=0
for i in range(2,j-2):
  soma2+=abs((4*jit[i]/5.-jit[i-1]/5.-jit[i-2]/5.-jit[i+1]/5.-jit[i+2]/5.))
soma2=soma2/(j-4)
ppq5=soma2/(sum(jit)/j)
ppq5=ppq5*100
soma3=0
for i in range(1,j-1):
  soma1+=abs((jit[i+1]+jit[i-1]-2*jit[i]))
soma3=soma3/(j-2)
ddp=soma3/(sum(jit)/j)
ddp=ddp*100
```

```
arquivo = open('parametros','w')
arquivo.write('jit = '+str(jitrel)+'% \n')
arquivo.write('rap = '+str(rap)+'% \n')
arquivo.write('ppq5 = '+str(ppq5)+'% \n')
arquivo.write('ddp = '+str(ddp)+'% \n')
arquivo.close()

#criacao do delta de dirac
imp = np.zeros(len(t))
imp[0] = 1
#convolucao entre o delta de dirac e o Pulso Glotal de Ros deterministico
z2 = np.convolve(g,imp)
#convolucao entre o delta de dirac e o Pulso Glotal de Ros com delta t(t) processo
    estocastico
z1 = np.convolve(x,imp)
#Trato Vocal
for H in range(1,6):
  L=['A','E','I','O','U']
  plt.subplot(2,1,1)
  plt.axis((0,3500*TO,0,Av+0.1))
  plt.plot(1000*t,z1[0:len(t)],'b')
  plt.ylabel('Amplitude')
  plt.title(u'EXCITACAO GLOTAL ANTES E DEPOIS DO FILTRO - VOGAL '+str(L[H-1]))

  #formantes da categoria de cantor baixo
  formants=[[900,1300,2000,2200,2500],
      [450,1700,2000,2200,2310],
      [300,1900,2100,2200,2490],
      [500,800,2150,2200,2490],
      [360,700,2170,2200,2330]]
  bandwidths=[41,52,70,32,100]
  fmts=formants[H-1]
  if FO>fmts[0]:
    fmts[0]=FO+10
  impulse = np.zeros(500)
  impulse[0] = 1
  xin=impulse
  for resonance in range(5):
    f=fmts[resonance]
    bw=bandwidths[resonance]
    num=np.array([1-2*exp(-bw*2*pi/fm)*cos(2*pi*f/fm)+exp(-4*bw*pi/fm)])
    den=np.array([1,-2*exp(-bw*2*pi/fm)*cos(2*pi*f/fm),exp(-4*bw*pi/fm)])
    yout=lfilter(num,den,xin,axis=0)
  xin=yout
  #Radiacao dos Labios (Aproximacao)
  r=0.95
  B1=np.array([1,-r])
  A1=np.array([1])
  yout=lfilter(B1,A1,np.squeeze(xin))
```

```
    #convolucao entre os pulso glotais e os filtros do trato vocal
    zout=np.convolve(yout,z1)
    ym=zout[0:len(x)]
    #salvar em áudio cada vogal sintetizada pelo modelo fonte-filtro
    sf.write('Ros_'+str(L[H-1])+'_a='+str(a)+'_c='+str(c)+'_eta='+str(eta)+'_jit'+str(
        jitrel)[0:5]+'%.wav', ym, fm)
    #concatena um trecho de cada vogal sustentada para gerar um áudio com todas vogais
    ym=ym[0:len(ym)/5]
    ym=np.concatenate((ym,np.zeros(len(ym))))
    todas=np.concatenate((todas,ym))

sf.write('Todas_Ros_a='+str(a)+'_c='+str(c)+'_eta='+str(eta)+'_jit'+str(jitrel)[0:5]+'%.
    wav', todas, fm) #salva o áudio de todas as vogais em um arquivo
```

Apêndice F

Pulso unitário determinístico de Liljencrants-Fant

```
# -*- coding: utf-8 -*-
import matplotlib.pyplot as plt
import numpy as np
from numpy import sin,cos,exp
from scipy.optimize import fsolve
from math import pi

#Pulso Unitario Deterministico de LF
def LF_uno_uni(Av,FO,fm,Oq,m,Qa):
  Tm=1/float(fm)
  TO=1/float(FO)
  T1=([])
  TOvetor=TO*np.ones((int(TO/Tm)+1,1)) #vetor com todos períodos instantâneos iguais a TO
  t1 = np.arange(0,Oq*TO,Tm) #tempo ate o ponto de inflexao
  t2 = np.arange(Oq*TO,TO,Tm) #tempo depois do ponto de inflexao
  t = np.arange(0, TO,Tm) #tempo total do ciclo
  x=([])

  #criacao do pulso de LF Deterministico
  for i in range(len(t1)):
    def f(E):
      return E*Qa*(1-Oq)*TOvetor[i]-1+exp(-E*(1-Oq)*TOvetor[i])
    E=fsolve(f,1000000)
    x0=pi/(m*Oq*TOvetor[i])
    x1=cos(pi/m)
    x2=sin(pi/m)
    x3=TOvetor[i]*(1-Oq)
    x4=exp(E*TOvetor[i]*(1-Oq))
    x5=1/E
    def g(a):
      return(1/(a**2+x0**2))*(a+x0*((exp(-a*Oq*TOvetor[i])-x1)/x2))-x3/(x4-1)+x5
```

```
        a=fsolve(g,0)
        x6=-Av*sin(pi/m)*(a**2+(pi/(m*Oq*TOvetor[i]))**2)/exp(-a*Oq*TOvetor[i])
        x7=pi/(m*Oq*TOvetor[i])+a*exp(a*Oq*TOvetor[i]*m)*sin(pi)-(pi/(m*Oq*TOvetor[i]))*exp(a
            *Oq*TOvetor[i]*m)*cos(pi)
        Ee=x6/x7
        x8=(-Ee*exp(-a*Oq*TOvetor[i])*((pi/(m*Oq*TOvetor[i]))+a*exp(a*t[i])*sin(pi*t[i]/(Oq*
            TOvetor[i]*m))-(pi/(m*Oq*TOvetor[i]))*exp(a*t[i])*cos(pi*t[i]/(Oq*TOvetor[i]*m)))
            )/((sin(pi/m)*(a**2+(pi/(m*Oq*TOvetor[i]))**2)))
        x=np.concatenate((x,x8))
    for i in range(len(t)-len(t1)):
        i=i+len(t1)
        def f(E):
            return E*Qa*(1-Oq)*TOvetor[i]-1+exp(-E*(1-Oq)*TOvetor[i])
        E=fsolve(f,1000000)
        x0=pi/(m*Oq*TOvetor[i])
        x1=cos(pi/m)
        x2=sin(pi/m)
        x3=TOvetor[i]*(1-Oq)
        x4=exp(E*TOvetor[i]*(1-Oq))
        x5=1/E
        def g(a):
            return(1/(a**2+x0**2))*(a+x0*((exp(-a*Oq*TOvetor[i])-x1)/x2))-x3/(x4-1)+x5
        a=fsolve(g,0)
        x6=-Av*sin(pi/m)*(a**2+(pi/(m*Oq*TOvetor[i]))**2)/exp(-a*Oq*TOvetor[i])
        x7=pi/(m*Oq*TOvetor[i])+a*exp(a*Oq*TOvetor[i]*m)*sin(pi)-(pi/(m*Oq*TOvetor[i]))*exp(a
            *Oq*TOvetor[i]*m)*cos(pi)
        Ee=x6/x7
        x9=-Ee*(1/(E*Qa*(1-Oq)*TOvetor[i])-1)*(TOvetor[i]-t[i]+(1-exp(E*(TOvetor[i]-t[i])))/E
            )
        x=np.concatenate((x,x9))

    imp=np.zeros(int(TO/Tm)+2)
    imp[0]=1
    z = np.convolve(x,imp)
    return x,t,TO
```

Apêndice G

Pulso unitário estocástico de Liljencrants-Fant e densidade espectral a 2 parâmetros

```
# -*- coding: utf-8 -*-
import matplotlib.pyplot as plt
import numpy as np
from numpy import sin,cos,exp
from scipy.optimize import fsolve
from math import pi

#Pulso Unitario com Processo Estocastico
def LF_uno_uni_normal(Av,FO,fm,Oq,m,Qa,abar,bbar):
  Tm=1/float(fm) #intervalo de amostragem do pulso
  TO=1/float(FO) #periodo fundamental do ciclo
  xbar=(1.-abar**2/(2*bbar))**0.5
  kbar=10000
  beg=3000
  xx=([])
  xini=0
  ko=1/40000.
  for ii in range(kbar):
    xfin=([((-float(bbar)*Tm/2.+1.)*xini+abar*np.random.randn()*Tm**0.5)/(1+float(bbar)*
        Tm/2.)])
    xx=np.concatenate((xx,xfin))
    xini=float(xfin[0])
  dT=ko+(Tm-ko)*(xbar+xx[beg:])**2
  TOfin=sum(dT[0:int(TO/Tm)])
  TOvetor=TOfin*np.ones(int(TO/Tm)+1)
  t=([0])
  for iii in range(len(dT)):
    t=np.concatenate((t,([t[-1]+dT[iii]])))
```

```
t=t[0:int(TO/Tm)+1]
x=([])
k=0
for i in range(len(t)):
    if t[i]<=Oq*TOfin:
        k+=1

#criacao do pulso de LF com delta t(t) processo estocástico
for i in range(k):
    def f(E):
        return E*Qa*(1-Oq)*TOvetor[i]-1+exp(-E*(1-Oq)*TOvetor[i])
    E=fsolve(f,1000000)
    x0=pi/(m*Oq*TOvetor[i])
    x1=cos(pi/m)
    x2=sin(pi/m)
    x3=TOvetor[i]*(1-Oq)
    x4=exp(E*TOvetor[i]*(1-Oq))
    x5=1/E
    def g(a):
        return(1/(a**2+x0**2))*(a+x0*((exp(-a*Oq*TOvetor[i])-x1)/x2))-x3/(x4-1)+x5
    a=fsolve(g,0)
    x6=-Av*sin(pi/m)*(a**2+(pi/(m*Oq*TOvetor[i]))**2)/exp(-a*Oq*TOvetor[i])
    x7=pi/(m*Oq*TOvetor[i])+a*exp(a*Oq*TOvetor[i]*m)*sin(pi)-(pi/(m*Oq*TOvetor[i]))*exp(a
        *Oq*TOvetor[i]*m)*cos(pi)
    Ee=x6/x7
    x8=(-Ee*exp(-a*Oq*TOvetor[i])*((pi/(m*Oq*TOvetor[i]))+a*exp(a*t[i])*sin(pi*t[i]/(Oq*
        TOvetor[i]*m))-(pi/(m*Oq*TOvetor[i]))*exp(a*t[i])*cos(pi*t[i]/(Oq*TOvetor[i]*m)))
        )/((sin(pi/m)*(a**2+(pi/(m*Oq*TOvetor[i]))**2)))
    x=np.concatenate((x,x8))
for i in range(k,len(t)):
    def f(E):
        return E*Qa*(1-Oq)*TOvetor[i]-1+exp(-E*(1-Oq)*TOvetor[i])
    E=fsolve(f,1000000)
    x0=pi/(m*Oq*TOvetor[i])
    x1=cos(pi/m)
    x2=sin(pi/m)
    x3=TOvetor[i]*(1-Oq)
    x4=exp(E*TOvetor[i]*(1-Oq))
    x5=1/E
    def g(a):
        return(1/(a**2+x0**2))*(a+x0*((exp(-a*Oq*TOvetor[i])-x1)/x2))-x3/(x4-1)+x5
    a=fsolve(g,0)
    x6=-Av*sin(pi/m)*(a**2+(pi/(m*Oq*TOvetor[i]))**2)/exp(-a*Oq*TOvetor[i])
    x7=pi/(m*Oq*TOvetor[i])+a*exp(a*Oq*TOvetor[i]*m)*sin(pi)-(pi/(m*Oq*TOvetor[i]))*exp(a
        *Oq*TOvetor[i]*m)*cos(pi)
    Ee=x6/x7
    x9=-Ee*(1/(E*Qa*(1-Oq)*TOvetor[i])-1)*(TOvetor[i]-t[i]+(1-exp(E*(TOvetor[i]-t[i])))/E
        )
    x=np.concatenate((x,x9))
```

```
    y = np.diff(x)/Tm #derivada do pulso no tempo
    imp=np.zeros(int(TOfin/Tm)+2) #função impulso delta(t)
    imp[0]=1
    z = np.convolve(x,imp) #convolucao entre o impulso unitario e o sinal do pulso que
        resulta no proprio sinal do pulso
    return x,t,TOfin
```

Apêndice H

Trem de pulsos estocásticos de Liljencrants-Fant e densidade espectral a 2 parâmetros - Cálculo das medidas de dispersão

```
import os
import matplotlib.pyplot as plt
import numpy as np
from numpy import cos,exp, log10,pi
from scipy.signal import lfilter, freqz
import soundfile as sf
from Pulso_Unitario_LF import LF_uno_uni_normal
from Pulso_Unitario_LF import LF_uno_uni
from random import random

a=14.1
b=100.0
#print os.getcwd()
os.makedirs('a='+str(a)[0:6]+'_b='+str(b))
os.chdir('C:/Users/diego/OneDrive/Documentos/Mestrado_UFF\Codigos\duasvariaveis\
    Codigos_1_duasvariaveis/a='+str(a)[0:6]+'_b='+str(b))
fm=20000 #frequencia de amostragem
Tm=1/float(fm) #periodo de amostragem
Av=7 #amplitude de vozeamento
Tfs=3 #tempo total de sustentacao da vogal
FO=98 #frequencia funcamental
TO=1/float(FO) #periodo fundamental
Oq=0.80 #quociente de abertura
m=0.75 #coeficiente de assimetria
Qa=0.3 #coeficiente da fase de retorno
```

```
x1=([]) #criacao do vetor de pulsos LF com delta t(t) processo estocastico
g1=([]) #criacao do vetor de pulsos LF Determinístico
jit=([])
todas=([])
t=([])
tz=0
#cria e concatena os pulsos
for i in range(int(Tfs/TO)+10):
  x,ty,TOfin=LF_uno_uni_normal(Av,FO,fm,Oq,m,Qa,a,b) #cria um pulso com delta t(t)
      processo estocastico
  x1=np.concatenate((x1,x))#concatena os pulsos
  tt=ty+tz
  t=np.concatenate((t,tt))
  tz=t[-1]
  jit=np.concatenate((jit,([TOfin])))
x=x1

for i in range(int(Tfs/TO)+1):
  g,t1,TOo=LF_uno_uni(Av,FO,fm,Oq,m,Qa) #cria um pulso LF determinístico
  g1=np.concatenate((g1,g)) #concatena os pulsos

g=g1
t1=np.arange(0,len(g)*Tm,Tm)

j=len(jit)
jitabs=0
for i in range(j-1):
  jitabs+=abs((jit[i]-jit[i+1]))
jitabs=jitabs/(j-1)
jitrel=jitabs/(sum(jit)/j)
jitrel=jitrel*100
print jitrel


#novos parâmetros
soma1=0
for i in range(1,j-1):
  soma1+=abs((2*jit[i]/3.-jit[i-1]/3.-jit[i+1]/3.))
soma1=soma1/(j-2)
rap=soma1/(sum(jit)/j)
rap=rap*100

soma2=0
for i in range(2,j-2):
  soma2+=abs((4*jit[i]/5.-jit[i-1]/5.-jit[i-2]/5.-jit[i+1]/5.-jit[i+2]/5.))
soma2=soma2/(j-4)
ppq5=soma2/(sum(jit)/j)
ppq5=ppq5*100
```

```
soma3=0
for i in range(1,j-1):
  soma1+=abs((jit[i+1]+jit[i-1]-2*jit[i]))
soma3=soma3/(j-2)
ddp=soma3/(sum(jit)/j)
ddp=ddp*100

arquivo = open('parametros','w')
arquivo.write('jit = '+str(jitrel)+'% \n')
arquivo.write('rap = '+str(rap)+'% \n')
arquivo.write('ppq5 = '+str(ppq5)+'% \n')
arquivo.write('ddp = '+str(ddp)+'% \n')
arquivo.close()

#criacao do delta de dirac
imp = np.zeros(len(t))
imp[0] = 1


#convolucao entre o delta de dirac e o Pulso Glotal de LF deterministico
z2 = np.convolve(g,imp)

#convolucao entre o delta de dirac e o Pulso Glotal de LF com deltat(t) processo
    estocastico
z1 = np.convolve(x,imp)

#Trato Vocal
for H in range(1,6):
  L=['A','E','I','O','U']
  plt.subplot(2,1,1)
  plt.axis((0,3500*TO,0,Av+0.1))
  plt.plot(1000*t,z1[0:len(t)],'b')
  plt.ylabel('Amplitude')
  plt.title(u'EXCITACAO GLOTAL ANTES E DEPOIS DO FILTRO - VOGAL '+str(L[H-1]))

  #formantes da categoria de cantor baixo
  formants=[[900,1300,2000,2200,2500],
       [450,1700,2000,2200,2310],
       [300,1900,2100,2200,2490],
       [500,800,2150,2200,2490],
       [360,700,2170,2200,2330]]
  bandwidths=[41,52,70,32,100]
  fmts=formants[H-1]
  if FO>fmts[0]:
    fmts[0]=FO+10
  impulse = np.zeros(500)
  impulse[0] = 1
  xin=impulse
```

```
    for resonance in range(5):
        f=fmts[resonance]
        bw=bandwidths[resonance]
        num=np.array([1-2*exp(-bw*2*pi/fm)*cos(2*pi*f/fm)+exp(-4*bw*pi/fm)])
        den=np.array([1,-2*exp(-bw*2*pi/fm)*cos(2*pi*f/fm),exp(-4*bw*pi/fm)])
        yout=lfilter(num,den,xin,axis=0)
        xin=yout

    #Radiacao dos Labios (Aproximacao)
    r=0.95
    B1=np.array([1,-r])
    A1=np.array([1])
    yout=lfilter(B1,A1,np.squeeze(xin))

    #convolucao entre os pulso glotais e os filtros do trato vocal
    zout=np.convolve(yout,z1)
    ym=zout[0:len(x)]

    #salvar em áudio cada vogal sintetizada pelo modelo fonte-filtro
    sf.write('LF_'+str(L[H-1])+'_a='+str(a)+'_b='+str(b)+'_jit'+str(jitrel)[0:5]+'%.wav',
        ym, fm)

    #concatena um trecho de cada vogal sustentada para gerar um áudio com todas vogais
    ym=ym[0:len(ym)/5]
    ym=np.concatenate((ym,np.zeros(len(ym))))
    todas=np.concatenate((todas,ym))
sf.write('Todas'+'_LF_a='+str(a)+'_b='+str(b)+'_jit'+str(jitrel)[0:5]+'%.wav', todas, fm)
    #salva o áudio de todas as vogais em um arquivo
```

Apêndice I

Pulso unitário estocástico de Liljencrants-Fant e densidade espectral a 3 parâmetros

```
# -*- coding: utf-8 -*-
import matplotlib.pyplot as plt
import numpy as np
from numpy import sin,cos,exp
from scipy.optimize import fsolve
from math import pi
from numpy.linalg import inv

#Pulso Unitario com Processo Estocastico
def LF_uno_uni_normal(Av,FO,fm,Oq,m,Qa,abar,bbar,eta):
  alpha=np.array([[0.,1.],[-float(bbar)**2,-2*eta*float(bbar)]])
  beta=np.array([[0],[abar]])
  Tm=1/float(fm) #intervalo de amostragem do pulso
  TO=1/float(FO) #periodo fundamental do ciclo
  xbar=(1.-abar**2/(4*eta*bbar**3))**0.5
  kbar=10000
  beg=3000
  xx=([])
  ko=1/40000.
  xini=np.array([[0],[0]])
  for ii in range(kbar):
    xfin=(-(alpha)*Tm/2.+np.ones((2,2))).dot(xini)
    xfin=(xfin+beta*np.random.randn()*Tm**0.5)
    xfin=inv((np.ones((2,2))+alpha*Tm/2.)).dot(xfin)
    kfin=([xfin[0][0]])
    xx=np.concatenate((xx,kfin))
    xini=xfin
  dT=ko+(Tm-ko)*(xbar+xx[beg:])**2
```

```
TOfin=sum(dT[0:int(TO/Tm)])
TOvetor=TOfin*np.ones(int(TO/Tm)+1)
t=([0])
for iii in range(len(dT)):
  t=np.concatenate((t,([t[-1]+dT[iii]])))
t=t[0:int(TO/Tm)+1]
x=([])
k=0
for i in range(len(t)):
  if t[i]<=Oq*TOfin:
    k+=1

#criacao do pulso de LF com delta t(t) processo estocástico
for i in range(k):
  def f(E):
    return E*Qa*(1-Oq)*TOvetor[i]-1+exp(-E*(1-Oq)*TOvetor[i])
  E=fsolve(f,1000000)
  x0=pi/(m*Oq*TOvetor[i])
  x1=cos(pi/m)
  x2=sin(pi/m)
  x3=TOvetor[i]*(1-Oq)
  x4=exp(E*TOvetor[i]*(1-Oq))
  x5=1/E
  def g(a):
    return(1/(a**2+x0**2))*(a+x0*((exp(-a*Oq*TOvetor[i])-x1)/x2))-x3/(x4-1)+x5
  a=fsolve(g,0)
  x6=-Av*sin(pi/m)*(a**2+(pi/(m*Oq*TOvetor[i]))**2)/exp(-a*Oq*TOvetor[i])
  x7=pi/(m*Oq*TOvetor[i])+a*exp(a*Oq*TOvetor[i]*m)*sin(pi)-(pi/(m*Oq*TOvetor[i]))*exp(a*Oq*TOvetor[i]*m)*cos(pi)
  Ee=x6/x7
  x8=(-Ee*exp(-a*Oq*TOvetor[i])*((pi/(m*Oq*TOvetor[i]))+a*exp(a*t[i])*sin(pi*t[i]/(Oq*TOvetor[i]*m))-(pi/(m*Oq*TOvetor[i]))*exp(a*t[i])*cos(pi*t[i]/(Oq*TOvetor[i]*m))))/((sin(pi/m)*(a**2+(pi/(m*Oq*TOvetor[i]))**2)))
  x=np.concatenate((x,x8))
for i in range(k,len(t)):
  def f(E):
    return E*Qa*(1-Oq)*TOvetor[i]-1+exp(-E*(1-Oq)*TOvetor[i])
  E=fsolve(f,1000000)
  x0=pi/(m*Oq*TOvetor[i])
  x1=cos(pi/m)
  x2=sin(pi/m)
  x3=TOvetor[i]*(1-Oq)
  x4=exp(E*TOvetor[i]*(1-Oq))
  x5=1/E
  def g(a):
    return(1/(a**2+x0**2))*(a+x0*((exp(-a*Oq*TOvetor[i])-x1)/x2))-x3/(x4-1)+x5
  a=fsolve(g,0)
  x6=-Av*sin(pi/m)*(a**2+(pi/(m*Oq*TOvetor[i]))**2)/exp(-a*Oq*TOvetor[i])
```

```
        x7=pi/(m*Oq*TOvetor[i])+a*exp(a*Oq*TOvetor[i]*m)*sin(pi)-(pi/(m*Oq*TOvetor[i]))*exp(a*Oq*TOvetor[i]*m)*cos(pi)
        Ee=x6/x7
        x9=-Ee*(1/(E*Qa*(1-Oq)*TOvetor[i])-1)*(TOvetor[i]-t[i]+(1-exp(E*(TOvetor[i]-t[i])))/E)
        x=np.concatenate((x,x9))

    y = np.diff(x)/Tm #derivada do pulso no tempo
    imp=np.zeros(int(TOfin/Tm)+2) #função impulso delta(t)
    imp[0]=1
    z = np.convolve(x,imp) #convolucao entre o impulso unitario e o sinal do pulso que resulta no proprio sinal do pulso
    return x,t,TOfin
```

Apêndice J

Trem de pulsos estocásticos de Liljencrants-Fant e densidade espectral a 3 parâmetros - Cálculo das medidas de dispersão

```
import os
import matplotlib.pyplot as plt
import numpy as np
from numpy import cos,exp, log10,pi
from scipy.signal import lfilter, freqz
import soundfile as sf
from Pulso_Unitario_LF import LF_uno_uni_normal
from Pulso_Unitario_LF import LF_uno_uni
from random import random

a=10.0
c=1.0
eta=0.1
#print os.getcwd()
os.makedirs('a='+str(a)[0:6]+'_c='+str(c)+'_eta='+str(eta))
os.chdir('C:/Users/diego/OneDrive/Documentos/Mestrado_UFF/Codigos/tresvariaveis/
    Codigos_1_tresvariaveis/a='+str(a)[0:6]+'_c='+str(c)+'_eta='+str(eta))
fm=20000 #frequencia de amostragem
Tm=1/float(fm) #periodo de amostragem
Av=7 #amplitude de vozeamento
Tfs=3 #tempo total de sustentacao da vogal
FO=98 #frequencia funcamental
b=2*pi*FO*c
TO=1/float(FO) #periodo fundamental
Oq=0.80 #quociente de abertura
```

```
m=0.75 #coeficiente de assimetria
Qa=0.3 #coeficiente da fase de retorno
x1=([]) #criacao do vetor de pulsos LF com delta t(t) processo estocastico
g1=([]) #criacao do vetor de pulsos LF Determinístico
jit=([])
todas=([])
t=([])
tz=0
#cria e concatena os pulsos
for i in range(int(Tfs/TO)+10):
  x,ty,TOfin=LF_uno_uni_normal(Av,FO,fm,Oq,m,Qa,a,b,eta) #cria um pulso com delta t(t)
      processo estocastico
  x1=np.concatenate((x1,x))#concatena os pulsos
  tt=ty+tz
  t=np.concatenate((t,tt))
  tz=t[-1]
  jit=np.concatenate((jit,([TOfin])))
x=x1
for i in range(int(Tfs/TO)+1):
  g,t1,TOo=LF_uno_uni(Av,FO,fm,Oq,m,Qa) #cria um pulso LF determinístico
  g1=np.concatenate((g1,g)) #concatena os pulsos
g=g1
t1=np.arange(0,len(g)*Tm,Tm)

j=len(jit)
jitabs=0
for i in range(j-1):
  jitabs+=abs((jit[i]-jit[i+1]))
jitabs=jitabs/(j-1)
jitrel=jitabs/(sum(jit)/j)
jitrel=jitrel*100
print jitrel

#novos parâmetros
soma1=0
for i in range(1,j-1):
  soma1+=abs((2*jit[i]/3.-jit[i-1]/3.-jit[i+1]/3.))
soma1=soma1/(j-2)
rap=soma1/(sum(jit)/j)
rap=rap*100
soma2=0
for i in range(2,j-2):
  soma2+=abs((4*jit[i]/5.-jit[i-1]/5.-jit[i-2]/5.-jit[i+1]/5.-jit[i+2]/5.))
soma2=soma2/(j-4)
ppq5=soma2/(sum(jit)/j)
ppq5=ppq5*100
soma3=0
for i in range(1,j-1):
  soma1+=abs((jit[i+1]+jit[i-1]-2*jit[i]))
```

```
soma3=soma3/(j-2)
ddp=soma3/(sum(jit)/j)
ddp=ddp*100

arquivo = open('parametros','w')
arquivo.write('jit = '+str(jitrel)+'% \n')
arquivo.write('rap = '+str(rap)+'% \n')
arquivo.write('ppq5 = '+str(ppq5)+'% \n')
arquivo.write('ddp = '+str(ddp)+'% \n')
arquivo.close()

#criacao do delta de dirac
imp = np.zeros(len(t))
imp[0] = 1


#convolucao entre o delta de dirac e o Pulso Glotal de LF deterministico
z2 = np.convolve(g,imp)

#convolucao entre o delta de dirac e o Pulso Glotal de LF com delta t(t) processo estocastico
z1 = np.convolve(x,imp)
plt.figure(figsize=(6.4,2.0))
plt.axis((0,1000*sum(jit[0:10]),0,Av+0.1))
plt.plot(1000*t[0:len(x)],z1[0:len(x)],'k',label=u'Delta t(t) Processo Estocástico')
plt.ylabel('Amplitude')
plt.xlabel('Tempo(ms)')
plt.title(u'Pulso Glotal')
plt.savefig('PULSOS_LF_'+'_a='+str(a)+'_c='+str(c)+'_jit'+str(jitrel)[0:5]+'%.png', format='png')
plt.show()


#Trato Vocal
for H in range(1,6):
  L=['A','E','I','O','U']
  plt.subplot(2,1,1)
  plt.axis((0,3500*TO,0,Av+0.1))
  plt.plot(1000*t,z1[0:len(t)],'b')
  plt.ylabel('Amplitude')
  plt.title(u'EXCITACAO GLOTAL ANTES E DEPOIS DO FILTRO - VOGAL '+str(L[H-1]))

  #formantes da categoria de cantor baixo
  formants=[[900,1300,2000,2200,2500],
      [450,1700,2000,2200,2310],
      [300,1900,2100,2200,2490],
      [500,800,2150,2200,2490],
      [360,700,2170,2200,2330]]
  bandwidths=[41,52,70,32,100]
```

```
    fmts=formants[H-1]
    if F0>fmts[0]:
        fmts[0]=F0+10
    impulse = np.zeros(500)
    impulse[0] = 1
    xin=impulse

    for resonance in range(5):
        f=fmts[resonance]
        bw=bandwidths[resonance]
        num=np.array([1-2*exp(-bw*2*pi/fm)*cos(2*pi*f/fm)+exp(-4*bw*pi/fm)])
        den=np.array([1,-2*exp(-bw*2*pi/fm)*cos(2*pi*f/fm),exp(-4*bw*pi/fm)])
        yout=lfilter(num,den,xin,axis=0)
        xin=yout

    #Radiacao dos Labios (Aproximacao)
    r=0.95
    B1=np.array([1,-r])
    A1=np.array([1])
    yout=lfilter(B1,A1,np.squeeze(xin))

    #convolucao entre os pulso glotais e os filtros do trato vocal
    zout=np.convolve(yout,z1)
    ym=zout[0:len(x)]

    #salvar em áudio cada vogal sintetizada pelo modelo fonte-filtro
    sf.write('LF_'+str(L[H-1])+'_a='+str(a)+'_c='+str(c)+'_eta='+str(eta)+'_jit'+str(jitrel)[0:5]+'%.wav', ym, fm)

    #concatena um trecho de cada vogal sustentada para gerar um áudio com todas vogais
    ym=ym[0:len(ym)/5]
    ym=np.concatenate((ym,np.zeros(len(ym))))
    todas=np.concatenate((todas,ym))
sf.write('Todas_LF_a='+str(a)+'_c='+str(c)+'_eta='+str(eta)+'_jit'+str(jitrel)[0:5]+'%.wav', todas, fm) #salva o áudio de todas as vogais em um arquivo
```

Apêndice K

Análise de convergência - Densidade espectral a 2 parâmetros

```python
# -*- coding: utf-8 -*-
import matplotlib.pyplot as plt
import numpy as np

Tm=1/20000.
abar=10.
bbar=10000.
xbar=(1.-abar**2/(2*bbar))**0.5
kbar=100000
k0=1/40000
xx=([])
dTbarra=([])
dTbarra2=([])
xini=0
n=1000
for ii in range(n):
    xfin=([((-float(bbar)*Tm/2.+1.)*xini+abar*np.random.randn()*Tm**0.5)/(1+float(bbar)*Tm
        /2.)])
    xx=np.concatenate((xx,xfin))
    xini=float(xfin[0])
    dT=ko+(Tm-ko)*(xbar+xx[beg:])**2
    barra=([sum(dT)/(ii+1.)])
    dT2=dT**2
    barra2=([sum(dT2)/(ii+1.)])
    dTbarra=np.concatenate((dTbarra,barra))
    dTbarra2=np.concatenate((dTbarra2,barra2))
dT=Tm*(xbar+xx)**2
dTquad=dT**2
N=np.arange(0,n,1)
plt.subplot(1,1,1)
plt.figure(figsize=(10,5))
```

```
plt.axis((0,n,0,1.5*Tm))
plt.plot(N,dTbarra,'k', linewidth=2.0)
plt.xlabel('Iterations')
plt.savefig('ConvergenciaT_t.png', format='png',bbox_inches='tight')
plt.show()
plt.subplot(1,1,1)
plt.figure(figsize=(10,5))
plt.axis((0,n,0,1.5*Tm**2))
plt.plot(N,dTbarra2,'k', linewidth=2.0)
plt.xlabel('Iterations')
plt.savefig('ConvergenciaT2_t.png', format='png',bbox_inches='tight')
plt.show()
```

Apêndice L

Função densidade de probabilidade - Densidade espectral a 2 parâmetros

```
# -*- coding: utf-8 -*-
import matplotlib.pyplot as plt
import numpy as np
from numpy import sin,cos,exp
from scipy.optimize import fsolve
from math import pi
from scipy.interpolate import spline

TO=1/98.
FO=1/float(TO)

Tm=1/20000.
abar=5.
bbar=7500.
xbar=(1.-abar**2/(2*bbar))**0.5
print xbar
kbar=5000
beg=1000
xx=([])
ko=1/40000
xini=0
TOs=([])
n=10000
for l in range(n):
for ii in range(beg+int(TO/Tm)+1):
xfin=([(((-float(bbar)*Tm/2.+1.)*xini+abar*np.random.randn()*Tm**0.5)/(1+float(bbar)*Tm/2.)])
xx=np.concatenate((xx,xfin))
xini=float(xfin[0])
dT=ko+(Tm-ko)*(xbar+xx[beg:])**2
TOfin=sum(dT[0:int(TO/Tm)])
```

```
TOs=np.concatenate((TOs,([TOfin])))
xx=([])
TOs=np.array(TOs)
FOs=1/TOs
delta=0.5
bins=np.arange(90,106,delta)
histogram=np.histogram(FOs,bins=bins,normed=True,density=True)[0]
bins=0.5*(bins[1:]+bins[:-1])
plt.figure(figsize=(9,6))
plt.xlabel('Frequency (Hz)')
#plt.plot(bins,histogram,'k',label='a='+str(abar)+'  b='+str(bbar))
## plt.show()

xnew = np.linspace(90,106,300) #300 represents number of points to make between T.min and
    T.max
power_smooth = spline(bins,histogram,xnew)
plt.plot(xnew,power_smooth,'k',label='a='+str(abar)+'  b='+str(bbar))
## plt.show()

Tm=1/20000.
##abar=10.
bbar=10000.
xbar=(1.-abar**2/(2*bbar))**0.5
print xbar
kbar=10000
beg=3000
xx=([])
xini=0
TOs=([])
for l in range(n):
  for ii in range(beg+int(TO/Tm)+1):
    xfin=([((-float(bbar)*Tm/2.+1.)*xini+abar*np.random.randn()*Tm**0.5)/(1+float(bbar)*
        Tm/2.)])
    xx=np.concatenate((xx,xfin))
    xini=float(xfin[0])
  dT=Tm*(xbar+xx[beg:])**2
  TOfin=sum(dT[0:int(TO/Tm)])
  TOs=np.concatenate((TOs,([TOfin])))
  xx=([])
TOs=np.array(TOs)
FOs=1/TOs
delta=0.5
bins=np.arange(90,106,delta)
histogram=np.histogram(FOs,bins=bins,normed=True,density=True)[0]
bins=0.5*(bins[1:]+bins[:-1])
##plt.plot(bins,histogram,'k:',label='a='+str(abar)+'  b='+str(bbar))
## plt.show()
```

```
xnew = np.linspace(90,106,300) #300 represents number of points to make between T.min and T.max
power_smooth = spline(bins,histogram,xnew)
plt.plot(xnew,power_smooth,'k:',label='a='+str(abar)+'  b='+str(bbar))
## plt.show()


Tm=1/20000.
##abar=10.
bbar=17500.
xbar=(1.-abar**2/(2*bbar))**0.5
print xbar
kbar=10000
beg=3000
xx=([])
xini=0
TOs=([])
for l in range(n):
  for ii in range(beg+int(TO/Tm)+1):
    xfin=([((-float(bbar)*Tm/2.+1.)*xini+abar*np.random.randn()*Tm**0.5)/(1+float(bbar)*Tm/2.)])
    xx=np.concatenate((xx,xfin))
    xini=float(xfin[0])
  dT=Tm*(xbar+xx[beg:])**2
  TOfin=sum(dT[0:int(TO/Tm)])
  TOs=np.concatenate((TOs,([TOfin])))
  xx=([])
TOs=np.array(TOs)
FOs=1/TOs
delta=0.5
bins=np.arange(90,106,delta)
histogram=np.histogram(FOs,bins=bins,normed=True,density=True)[0]
bins=0.5*(bins[1:]+bins[:-1])
# plt.plot(bins,histogram,'k--',label='a='+str(abar)+'  b='+str(bbar))
##plt.show()
xnew = np.linspace(90,106,300) #300 represents number of points to make between T.min and T.max
power_smooth = spline(bins,histogram,xnew)
plt.plot(xnew,power_smooth,'k--',label='a='+str(abar)+'  b='+str(bbar))
plt.legend(loc=0)
plt.savefig(u'variando_b_'+str(FO)+'Hz_a='+str(abar)+'.png', format='png',bbox_inches='tight')
plt.show()
```

Apêndice M

Função densidade espectral de potência a 3 parâmetros

```
#!/usr/bin/env python
# -*- coding: utf-8 -*-
import matplotlib.pyplot as plt
import numpy as np
from math import pi
%from numpy.linalg import inv

freq=180.
a=10.
b=2*pi*freq
eta=0.1
k=2*freq
f=np.arange(0,k,1)
s=(1./(2*pi))*a**2/((b**2-(2*pi*f)**2)**2+4*eta**2*b**2*(2*pi*f)**2)
smax=max(s)
s=s/smax
plt.figure(figsize=(10,5))
plt.axis((0,k,0,1.1*max(s)))
plt.plot(f,s,'k',label=r'$\xi$= '+str(eta))
eta=0.15
s=(1./(2*pi))*a**2/((b**2-(2*pi*f)**2)**2+4*eta**2*b**2*(2*pi*f)**2)
s=s/smax
plt.plot(f,s,'k--',label=r'$\xi$= '+str(eta))
eta=0.30
s=(1./(2*pi))*a**2/((b**2-(2*pi*f)**2)**2+4*eta**2*b**2*(2*pi*f)**2)
s=s/smax
plt.plot(f,s,'k:',label=r'$\xi$= '+str(eta))
eta=0.70
s=(1./(2*pi))*a**2/((b**2-(2*pi*f)**2)**2+4*eta**2*b**2*(2*pi*f)**2)
s=s/smax
plt.plot(f,s,'k-.',label=r'$\xi$= '+str(eta))
```

```
plt.xlabel('Frequency (Hz)')
plt.legend(loc=0)
plt.savefig(u'variando_eta.png', format='png',bbox_inches='tight')
plt.show()

freq=180.
a=10.
c=1.0
b=2*pi*freq*c
eta=0.2
k=3*freq
f=np.arange(0,k,1)
s=(1./(2*pi))*a**2/((b**2-(2*pi*f)**2)**2+4*eta**2*b**2*(2*pi*f)**2)
smax=max(s)
s=s/smax
plt.figure(figsize=(10,5))
plt.axis((0,k,0,1.1*max(s)))
plt.plot(f,s,'k',label=r'$c_b$= '+str(c))
c=1.2
b=2*pi*freq*c
s=(1./(2*pi))*a**2/((b**2-(2*pi*f)**2)**2+4*eta**2*b**2*(2*pi*f)**2)
s=s/smax
plt.plot(f,s,'k--',label=r'$c_b$= '+str(c))
c=1.5
b=2*pi*freq*c
s=(1./(2*pi))*a**2/((b**2-(2*pi*f)**2)**2+4*eta**2*b**2*(2*pi*f)**2)
s=s/smax
plt.plot(f,s,'k:',label=r'$c_b$= '+str(c))
c=2.0
b=2*pi*freq*c
s=(1./(2*pi))*a**2/((b**2-(2*pi*f)**2)**2+4*eta**2*b**2*(2*pi*f)**2)
s=s/smax
plt.plot(f,s,'k-.',label=r'$c_b$= '+str(c))
plt.xlabel('Frequency (Hz)')
plt.legend(loc=0)
plt.savefig(u'variando_c.png', format='png',bbox_inches='tight')
plt.show()
```