

# Implementação da síntese FM em uma linha de atraso variável e suas possíveis aplicações no processamento de áudio em tempo real

Sérgio Freire

Escola de Música da UFMG - Campus Pampulha - Av. Antônio Carlos, 6627  
31270-010 – Belo Horizonte – MG – Brasil

sfreire@musica.ufmg.br

***Resumo:** O artigo apresenta uma proposta de transposição das técnicas e do repertório de timbres da síntese FM para o processamento digital de sons quaisquer, baseada no uso de linhas de atraso variáveis. Os cálculos necessários para uma correta adaptação dos parâmetros básicos da FM a essas linhas de atraso são acompanhados da discussão das possibilidades e de alguns cuidados em sua aplicação a sons instrumentais.*

***Abstract:** The paper proposes a transposition of the techniques and the timbral repertoire developed by the FM synthesis to the digital processing of any sounds, which is based on variable delay lines. The calculation of the parametrical equivalence between both techniques is followed by a discussion of the possibilities – and some cautions – to be considered in its use with instrumental sounds.*

## Introdução

O presente trabalho trata das possibilidades de aplicação da grande experiência adquirida com a síntese digital de som por meio da modulação de frequência (FM) no processamento de sinais de áudio em tempo real. Sua fundamentação se baseia no fato de que o efeito doppler pode ser visto como um caso particular de modulação em frequência (FM). O efeito doppler é facilmente obtido com uma das mais antigas técnicas de processamento sonoro – o atraso (*delay*) -, cuja versão digital oferece considerável flexibilidade na construção de linhas de atraso variável. Propõe-se aqui uma implementação da síntese FM baseada em uma linha de atraso variável equivalente à clássica implementação descrita por John Chowning em 1973. Em seguida são discutidas algumas particularidades e cuidados que devem ser tomados tanto em seu uso com um sinal de áudio em tempo real quanto na aplicação de formas de onda moduladoras diferentes da senoide.

## Síntese FM e efeito doppler

A modulação em frequência tem por base a modificação da frequência de um sinal senoidal (chamada de frequência portadora) por um outro sinal senoidal (frequência moduladora). Ao se elevar a frequência da moduladora para uma ordem de grandeza auditiva (acima de 16 Hz), acontecem interações capazes de gerar uma variada gama de timbres. Os parâmetros de um sinal modulado em frequência são os seguintes: frequência portadora, frequência moduladora e desvio máximo da frequência portadora. A amplitude instantânea de uma onda modulada em frequência é dada pela fórmula:

$e = A \sin(2\pi f_p t + I \sin 2\pi f_m t)$ , onde  $A$  é a amplitude máxima da modulação resultante,  $f_p$  é a frequência (em Hz) da onda portadora,  $f_m$  é frequência da moduladora e  $I$  é o índice de modulação, definido pela razão entre o desvio máximo (da portadora) e a frequência da onda moduladora<sup>1</sup>. A variação temporal do índice de modulação permite a construção relativamente simples de timbres que apresentam uma evolução espectral característica. Outro fator de grande relevância sobre o resultado sonoro é a razão existente entre as frequências moduladora e portadora, responsável pelo grau de harmonicidade do timbre resultante.

O efeito doppler é percebido, no caso da audição, como uma alteração da frequência do som escutado, sendo causado pela variação da distância existente entre a fonte sonora e o ouvinte. Ou seja, o emissor e o receptor sonoros não mantêm entre si uma distância constante, que pode ser expressa por uma velocidade fixa ou variável no tempo. O cálculo da frequência resultante é expresso pela seguinte fórmula:  $f = f_0 * \left(1 + \frac{v}{v_s}\right)$ , onde  $f_0$  é frequência emitida pela fonte sonora,  $v$  a velocidade existente entre o ouvinte e a fonte, e  $v_s$  a velocidade do som no ar.

### Cálculo dos parâmetros da linha de atraso

Pode-se interpretar o atraso provocado em um sinal digital de áudio como a criação de uma distância entre este sinal e o ouvinte. Cada amostra representa uma duração de  $(1/f_a)$ , onde  $f_a$  é a frequência de amostragem. Em uma situação física, uma atraso de  $N$  amostras equivale a uma distância de  $((N/f_a) * v_s)$  da fonte em relação ao ouvinte. Variar esse atraso significa imprimir uma velocidade relativa entre a fonte e o ouvinte, causando, por consequência, uma variação de frequência.

A implementação aqui proposta se baseia nos mesmos parâmetros da síntese FM, com a diferença de que o desvio máximo deve ser agora medido em amostras. Para se encontrar a relação existente entre os parâmetros de cada caso é necessário primeiramente transformar a variação do número de amostras (respectivamente da distância) em uma variação de velocidade, o que é facilmente obtido pela derivação da equação que representa o desvio instantâneo. Assim, se

$\Delta n(t) = -\frac{N_{\max}}{f_a} * v_s * \cos 2\pi f_m t$  representa a variação instantânea da distância entre a fonte e o ouvinte<sup>2</sup>,

$\Delta v(t) = \frac{\partial \Delta n(t)}{\partial t} = \frac{N_{\max}}{f_a} * v_s * 2\pi f_m * \sin 2\pi f_m t$  representa a velocidade instantânea entre eles.

Com isso, torna-se possível uma comparação direta entre as variações de frequência em cada um dos casos. A equação seguinte tem como seu membro esquerdo

<sup>1</sup> Ver Chowning (1973), p. 527. Ver também Tempelaars (1996), p. 248-253, para uma descrição mais detalhada das equações envolvidas na FM.

<sup>2</sup> A opção pela função cosseno com sinal negativo representa uma escolha da fase inicial da variação do atraso. Essa escolha coloca a variação de frequência na linha de atraso em fase com a variação presente na FM, já que a derivada de  $-\cos(x)$  equivale a  $\sin(x)$ .

a variação de frequência devida ao efeito doppler, e no direito a variação de frequência em uma síntese FM:

$$f_p \left( 1 + \frac{\frac{N_{\max} * v_s * 2\pi f_m * \sin 2\pi f_m t}{f_a}}{v_s} \right) = f_p + D_{\max} * \sin 2\pi f_m$$

Desenvolvendo ambos os termos e realizando as simplificações possíveis, obtém-se:

$$\frac{N_{\max}}{f_a} = \frac{D_{\max}}{2\pi f_p f_m}. \text{ Como } D_{\max} \text{ é definido pelo produto de } I \text{ por } f_m, \text{ chega-se a:}$$

$$\frac{N_{\max}}{f_a} = \frac{I}{2\pi f_p}. \quad (1)$$

O primeiro termo dessa equação é o desvio máximo medido em segundos, que depende – de modo diretamente proporcional – do índice de modulação e – de modo inversamente proporcional – da frequência portadora.

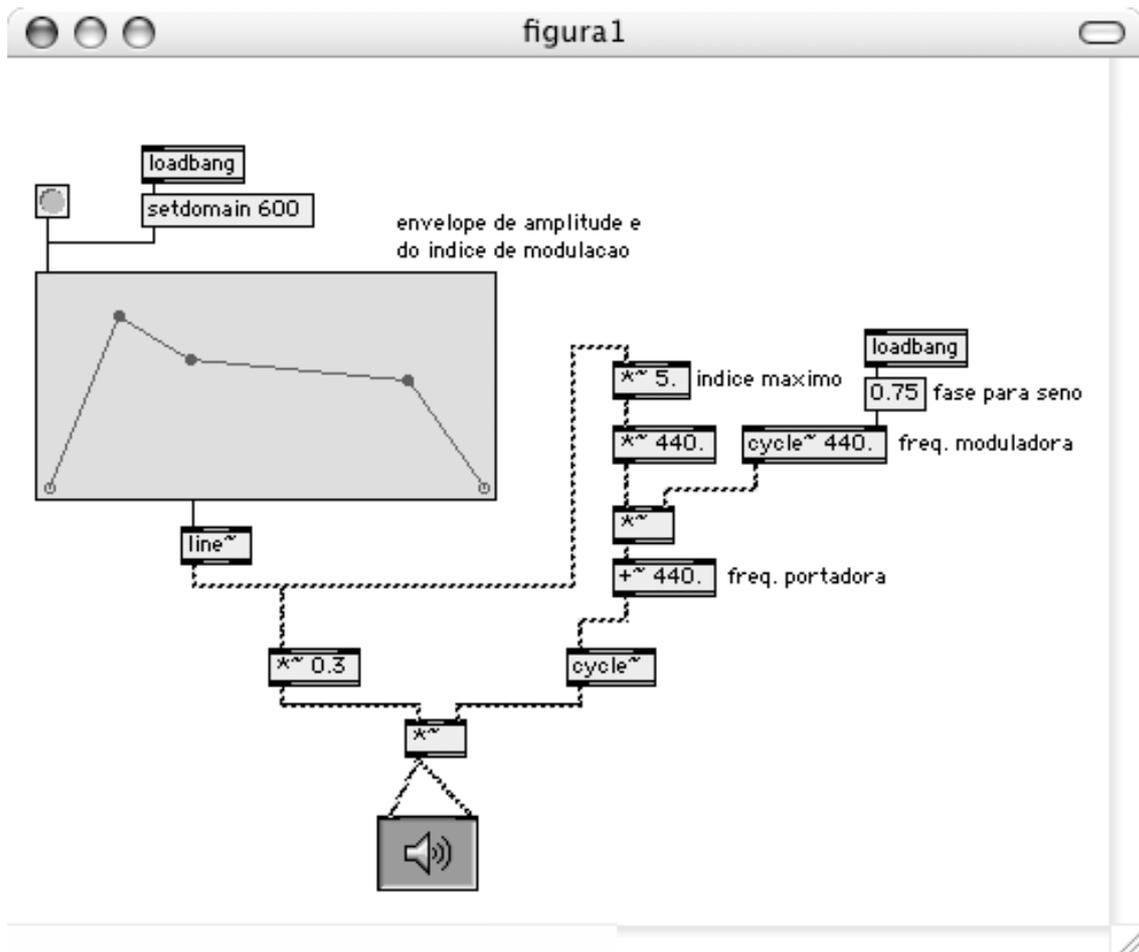
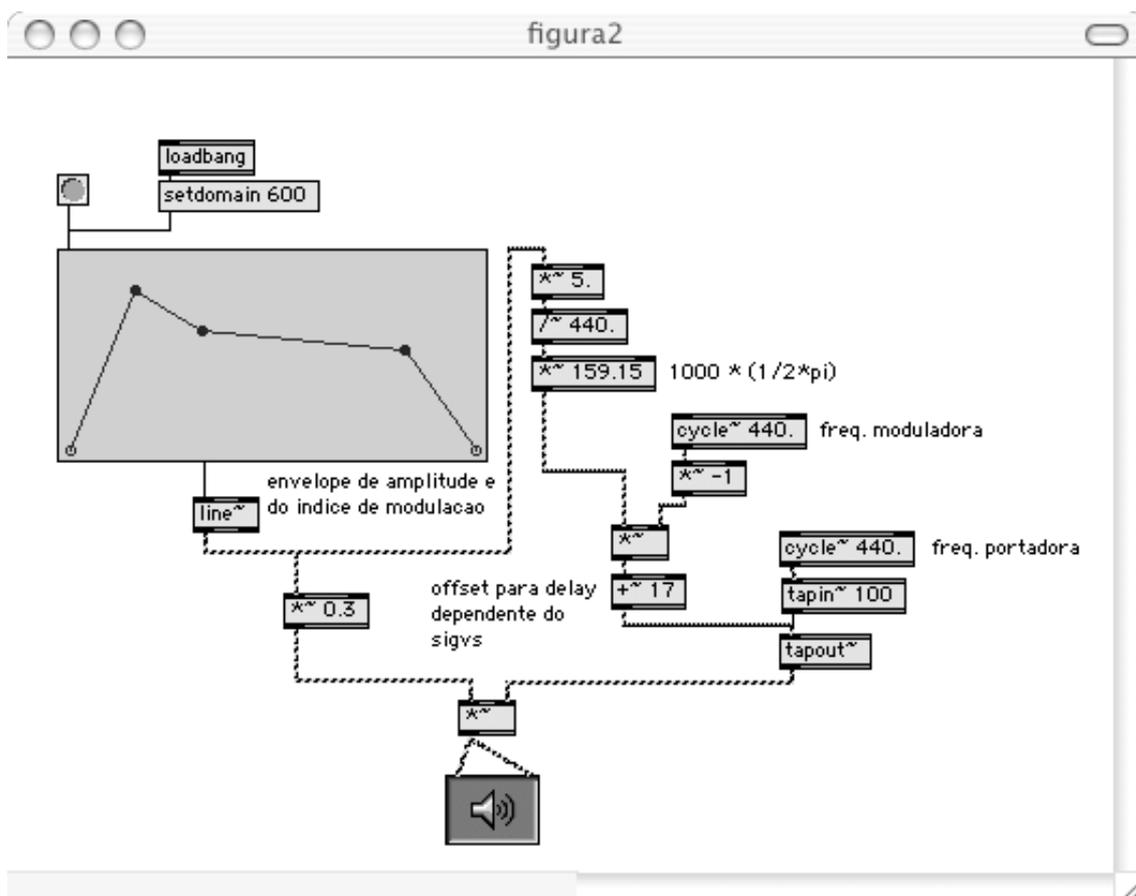


Figura 1: Implementação em max-msp da modulação em frequência de *brass-like tones* – descrita por Chowning (1973), p. 532.



**Figura 2: Implementação em max-msp da mesma modulação da figura 1, usando uma linha de atraso variável.**

As figuras acima representam a implementação (na linguagem *max-msp*) de uma modulação em frequência de um sinal senoidal: a primeira se utiliza de osciladores do tipo *look-up table*, a segunda faz uso de uma linha de atraso variável. O exemplo se refere à simulação de instrumentos de metal descrita por Chowning (1973), na qual a duração do som resultante é de 0.6 segundos, as frequências portadora e moduladora têm o valor 440 Hz, e o envelope - definido entre os valores limites 0 e 1 - se aplica tanto à amplitude (modificado por um fator de amplificação / atenuação) quanto ao índice de modulação (multiplicado por 5). Note-se que a unidade temporal usada é o milissegundo, o que faz com que as durações e o desvio máximo do atraso estejam multiplicados por 1000.

Essa implementação está baseada em alguns pressupostos: (a) o sinal de áudio que alimenta a linha de atraso deve ser invariável no tempo, tornando-se, assim, equivalente a um oscilador do tipo *look-up table*; (b) o tamanho do *buffer* selecionado deve ser capaz de atender à variação máxima de atraso, tanto no sentido positivo quanto negativo; ou, alternativamente, deve-se implementar um mecanismo de *wrapping* dos valores de atraso para que caiam sempre dentro dos limites do *buffer*, cujo módulo seja múltiplo da frequência portadora. A título de exemplo: a modulação de uma onda portadora de 20 Hz com um índice máximo de modulação de 25 gera um atraso máximo

de ca. 199 ms (cálculo advindo da equação 1), o que requer, sem o mecanismo de *wrapping*, um *buffer* mínimo de 398 ms.<sup>3</sup>

### **Modulação em frequência de sinais de áudio: precauções e possibilidades**

Com o exposto acima, está aberto o caminho para a modulação em frequência de sinais de áudio quaisquer: basta conectá-los à entrada da linha de atraso. Mas, como utilizar esse recurso de processamento, mantendo parâmetros mínimos de comparação e alguma previsibilidade nessa experimentação? Minha primeira opção é a de limitar seu uso a aplicações em tempo real e a sons gerados por um único instrumento: aliam-se aqui uma fácil implementação e reais possibilidades de extensão tímbrica instrumental.

Nesse campo de exploração, os seguintes pontos devem ser considerados, levando-se em conta ainda que diferentes graus de independência ou hierarquia podem aparecer entre eles, dependendo de contextos específicos:

a) os sons a serem modulados em frequência apresentam, obviamente, um espectro harmônico mais rico do que o da onda portadora original (senóide). Esses harmônicos também sofrem a modulação em frequência, e respondem ao índice de modulação de modo proporcional ao seu índice na série harmônica. Ou seja, o segundo harmônico (cuja frequência é o triplo da frequência fundamental) sofrerá uma modulação efetiva três vezes mais forte do que o som fundamental<sup>4</sup>. Sons percussivos com espectro inarmônico podem ganhar uma nova configuração energética entre seus parciais sem perder totalmente sua identidade. Deve-se contar ainda com a possibilidade de surgimento de *aliasing*; para evitá-lo, basta filtrar o sinal de entrada, sempre que necessário.

b) sons instrumentais, são, via de regra, variáveis no tempo, o que gera automaticamente um espectro com características dinâmicas. Porém, variações muito bruscas na forma de onda de entrada (principalmente em frequência e timbre) podem criar modulações inesperadas ou não desejadas.

c) um controle mais refinado do índice de harmonicidade requer o conhecimento preciso da frequência do sinal de entrada. Um algoritmo de detecção de alturas pode ser usado<sup>5</sup>, ou mesmo uma partitura com clara definição das alturas a serem tocadas.

d) é importante definir como será feito o controle do envelope dinâmico do processamento. Na síntese FM, o disparo de um novo som inicia automaticamente os envelopes relativos à amplitude e ao índice de modulação. Aqui, diferentes situações podem ser imaginadas, que vão do simples disparo de envelopes pré-fixados ao controle exercido pelo próprio envelope do som de entrada. Outras possibilidades são abertas pelo uso de controladores contínuos (um pedal de volume, p. ex) por parte do músico.

Pode-se também imaginar que um som instrumental de curta duração seja capaz de gerar um evento mais longo: nesse caso, deve-se implementar algum tipo de

---

<sup>3</sup> A implementação de um oscilador do tipo *look-up table* baseia-se fundamentalmente nessa técnica de *wrapping*, já que pressupõe que um ciclo da onda a represente em uma duração infinita. Moore (1992), p. 159-173, apresenta uma discussão detalhada da implementação de um oscilador desse tipo.

<sup>4</sup> O mesmo se dá em relação a uma onda complexa que serve a um oscilador do tipo *look-up table*. Um ciclo completo do oscilador gera um ciclo para a frequência fundamental, ao mesmo tempo em que gera três ciclos para o segundo harmônico.

<sup>5</sup> Tais como *fiddle~*, de Miller Puckette, ou *pitch~*, de Tristan Jehan.

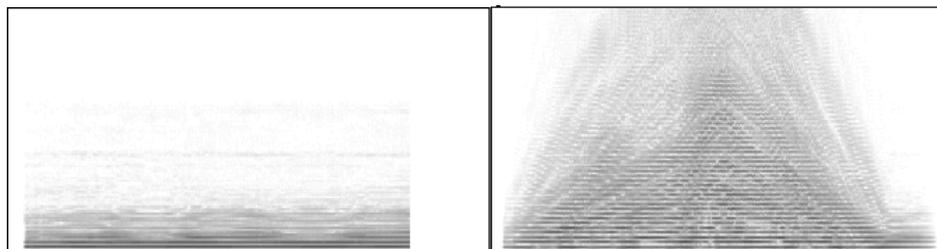
realimentação da linha de atraso que, como no caso do *wrapping* descrito mais acima, deve manter uma relação de números inteiros com o período do som de entrada (onda portadora).

e) os envelopes de amplitude, do índice de modulação e da harmonicidade podem ser controlados de forma totalmente independente; o contexto de sua aplicação é que deve determinar o grau de interdependência.

f) formas de onda alternativas à senóide podem ser usadas para a modulação. Nessa situação, a experiência adquirida com a síntese FM pode ser de grande utilidade para uma certa previsibilidade dos resultados. De maneira bem geral, pode-se afirmar que a influência de uma certa forma de onda sobre a modulação do atraso é equivalente ao efeito de sua derivada aplicada à forma de onda da moduladora em uma síntese FM. Assim, uma onda triangular aplicada à modulação do atraso corresponde a uma onda quadrada aplicada à modulação da frequência de uma onda portadora; o mesmo se dá entre uma onda quadrada e um trem de pulsos.

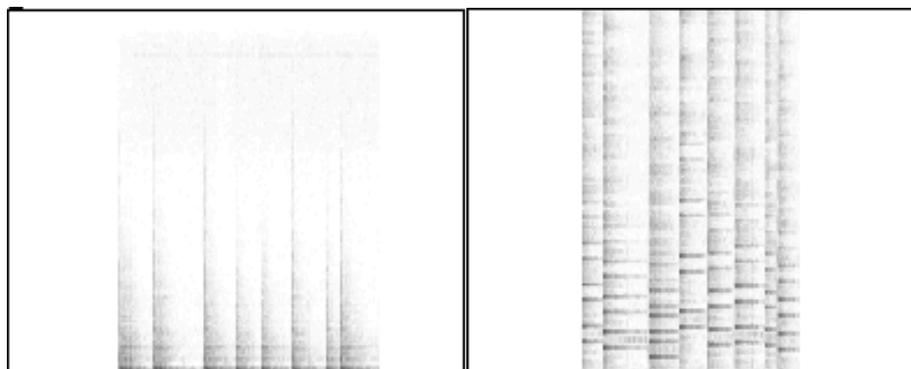
Dois exemplos bastante simples podem ajudar a ilustrar algumas possibilidades dessa modalidade de processamento:

(1) uma nota sustentada de clarineta é modulada em frequência, com índice de harmonicidade igual a 2, e índice de modulação variando de 0 a 5, e novamente até 0. A figura representa o espectrograma do som original e do som modulado.



**Figura 3: sonogramas de uma nota longa de clarineta (à esquerda), e desta mesma nota modulada em frequência (à direita).**

(2) um som de tamborim repetido algumas vezes é modulado com diferentes graus de harmonicidade, com um índice de modulação igual a 2.5 (é utilizada uma frequência fundamental de 481.7 Hz). O novo timbre apresenta, ao lado da variação em alturas, uma acentuada característica metálica.



**Figura 4: sonogramas de uma seqüência rítmica de um tamborim (à esquerda), e desta mesma seqüência modulada em frequência (à direita).**

## **Conclusão**

A implementação proposta, embora extremamente simples, apresenta um grau considerável de flexibilidade para aplicações em tempo real, principalmente na extensão dos timbres de instrumentos tradicionais. Boa parte dessa flexibilidade vem da impossibilidade de sua aplicação imediata, já que são necessárias diversas definições adicionais: detecção da frequência portadora, tipo de controle dos diferentes envelopes, espectros harmônicos esperados etc. Com isso, espera-se que boa parte da riqueza de timbres, que deu à síntese FM uma relativa hegemonia na geração sonora nos anos 1980 e 90, possa ser adaptada ao processamento de sons de instrumentos acústicos.

## **Bibliografia**

Chowning, John M. (1973), "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation", *Journal of the Audio Engineering Society* 21(7): 526-534.

Moore, Richard F., *Elements of Computer Music*, Englewood Cliffs: Prentice-Hall, 1992.

Puckette, Miller, Apel T. (1998), "Real-time audio analysis tools for PD and MSP", *Proceedings of the ICMC 1998*. San Francisco, Internacional Computer Music Association, p. 109-112.

Schottstaedt, B. (1977), "The Simulation of Natural Instrument Tones Using Frequency Modulation with a Complex Modulating Wave", *Computer Music Journal* 1(4):46-50.

Tempelaars, Stan, *Signal Processing, Speech and Music*, Lisse: Swets & Zeitlinger, 1996.