

## SÍNTESIS DE SONIDO Y EFECTOS DE AUDIO DIGITAL

Isabel Barbancho  
Profesora Titular de Universidad  
Universidad de Málaga  
E.T.S. Ingeniería de Telecomunicación  
Campus de Teatinos s/n, 29071 Málaga  
Teléfono: 952132587  
email: [ibp@ic.uma.es](mailto:ibp@ic.uma.es)

Noviembre, 2011

1

### ***Organización de la charla***

---

1. Introducción
2. Técnicas de síntesis de sonido
3. Efectos de sonido
4. Conclusiones

2

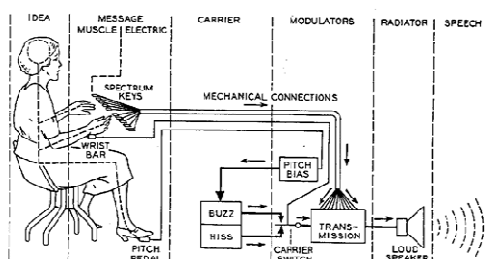
## Introducción

- Qué se entiende por síntesis de sonido?
- Breve historia
- Notas musicales
- Sonido
- Percepción del sonido
- Características temporales
- Características frecuenciales

3

## Introducción

- **¿Qué se entiende por síntesis de sonido?**
  - Técnicas de procesado de señal aplicadas a replicar sonidos o crear nuevos sonidos.
- **Historia de la síntesis de sonido**
  - Primer ejemplo: Vocoder 1939. Homer Dudley, Bell Laboratories (New Jersey). Tecnología de válvulas



4

## Introducción

- **Historia: Síntesis analógica**

- Uno de los primeros en sintetizar música. Sintetizador de Moog. Principios de los '50. Tecnología de transistores.

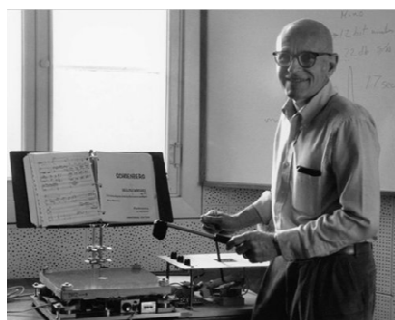


5

## Introducción

- **Historia: Síntesis Digital**


- Max Mathews (Bell Lab) en los '60 → generación de sonidos por ordenador.
  - Padre de la Música por ordenador y de las DIGITAL MUSIC TOOLS
  - MUSIC, primer programa ampliamente empleado para generación de sonido (1957)
  - No se podía usar en tiempo real → ordenadores muy lentos



6

## ***Introducción***

---

- **Historia: Síntesis Digital**
  - John Chowning en 1966 → síntesis FM → tiempo real asequible
  -  YAMAHA realizó el chip DX-7 en los 70



7

## ***Introducción***

---

- **Historia: Síntesis Digital**
  - Julius Smith en los '80 → síntesis de sonido mediante modelado por guía de onda
  - En los '90 síntesis wavetable

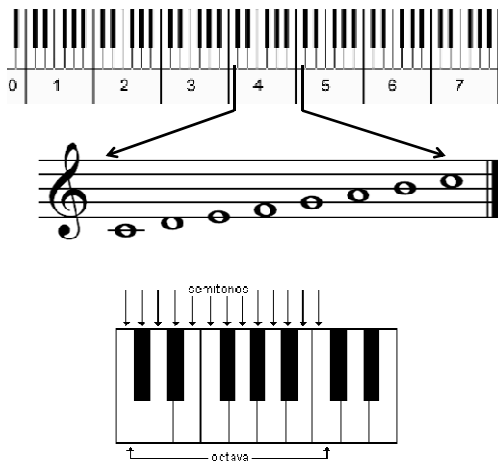
8

### Notas musicales

- Cada nota musical lleva asociada una determinada frecuencia fundamental ('pitch')

- Escala temperada

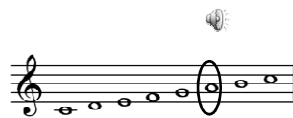
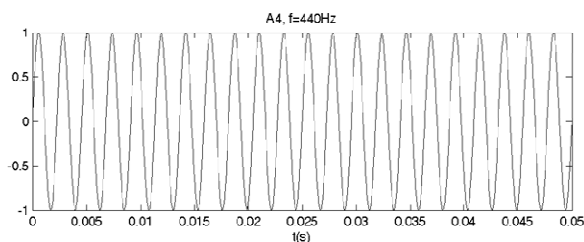
	Nota	Frecuencia (pitch)	Núm. MIDI
Do4	C4	261.63 Hz	60
Re4	D4	293.66Hz	62
Mi4	E4	329.63Hz	64
Fa4	F4	349.23Hz	65
Sol4	G4	392.00Hz	67
La4	A4	440.00Hz	69
Si4	B4	493.88Hz	71
Do5	C5	493.88Hz	72



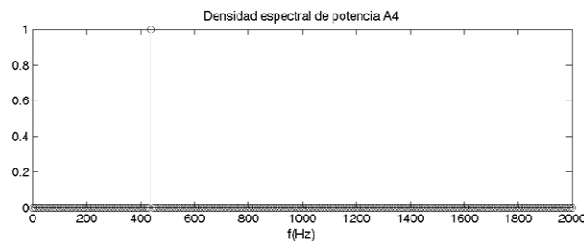
9

### Notas musicales

- Forma de onda en tiempo y espectro de la nota A4



Nota	Frecuencia	Núm. MIDI
A4	440.00Hz	69



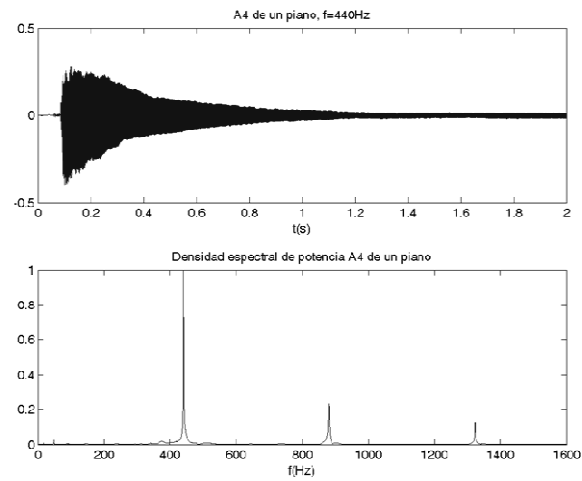
$$y = A \sin(2\pi f t)$$

```
fs=44100;
f=440;
t=0:1/fs:1;
y=sin(2*pi*f*t);
sound(y,fs)
```

10

## Notas musicales

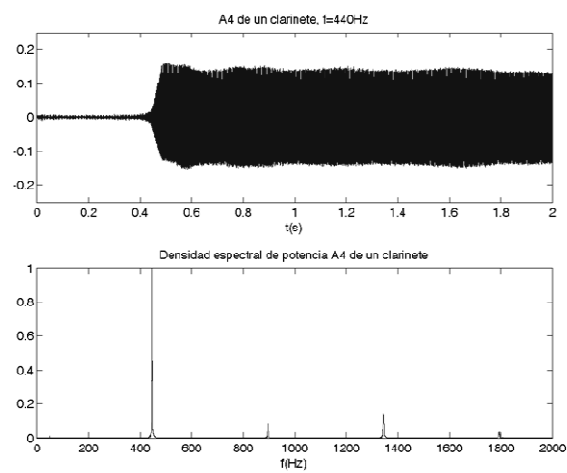
- Nota A4 de un piano



11

## Notas musicales

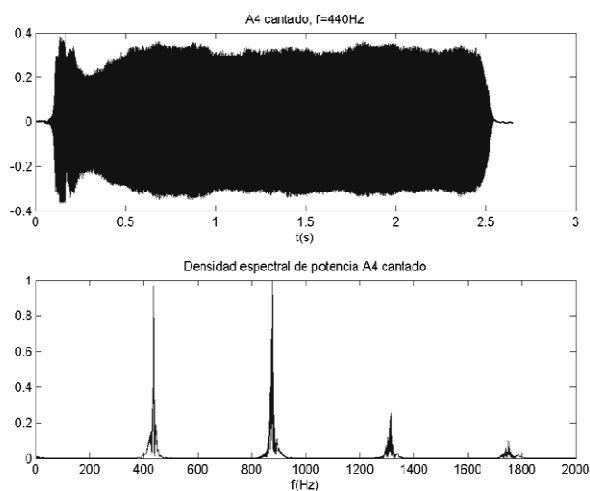
- Nota A4 de un clarinete



12

## Notas musicales

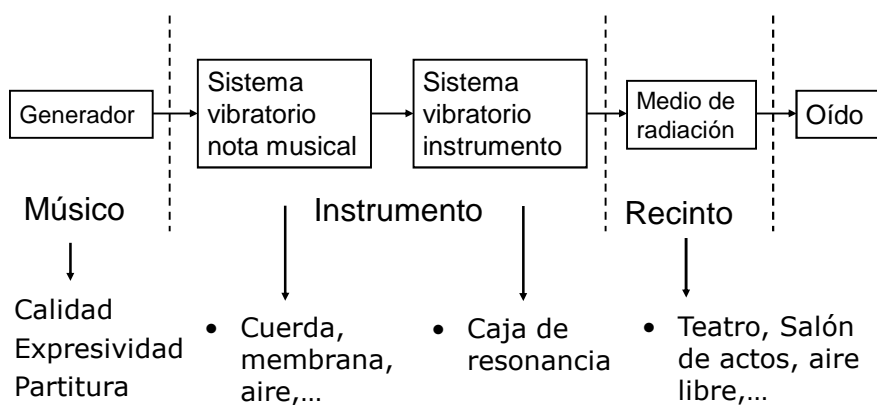
- Nota A4 voz



13

## Sonido

- Sistema de generación/recepción de sonido

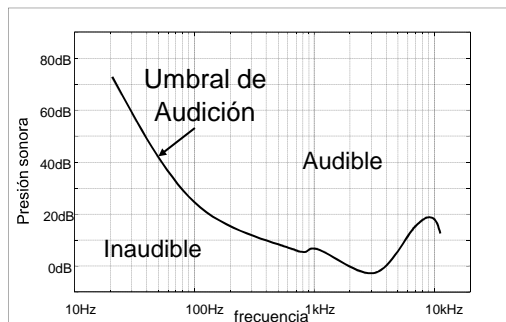


14

## Percepción del sonido

- Características del sistema auditivo

- Rango audible: (20Hz-20kHz)
- No presenta una respuesta plana
- Bastante insensibilidad a la fase del espectro de los sonidos

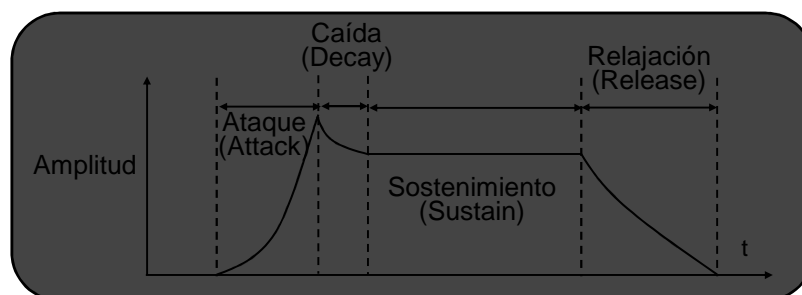


Influye en el diseño de las técnicas de síntesis musicales

15

## Características temporales

- La envolvente de la señal varía con el tiempo (ADSR)

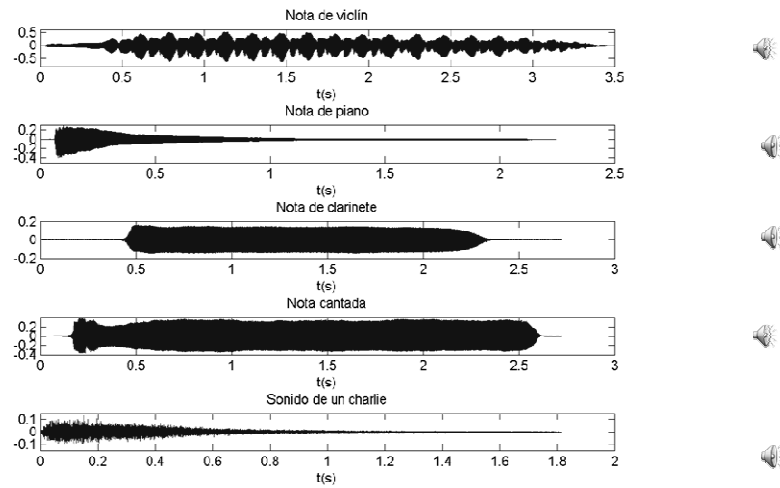


¡¡Los tiempos de ataque y relajación son los más determinantes en la identificación del instrumento!!

16



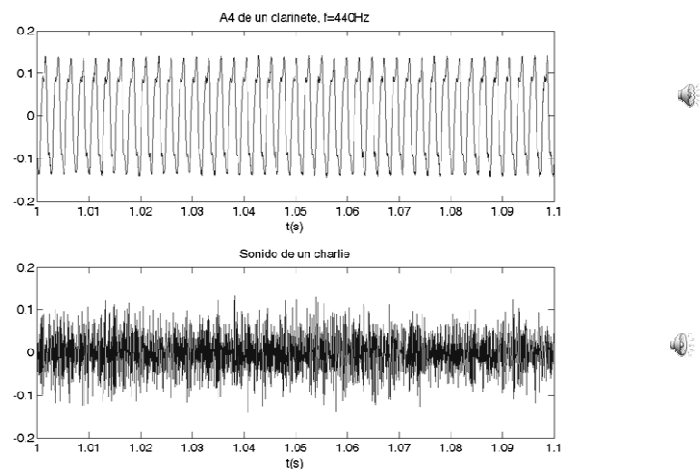
## Características temporales



17

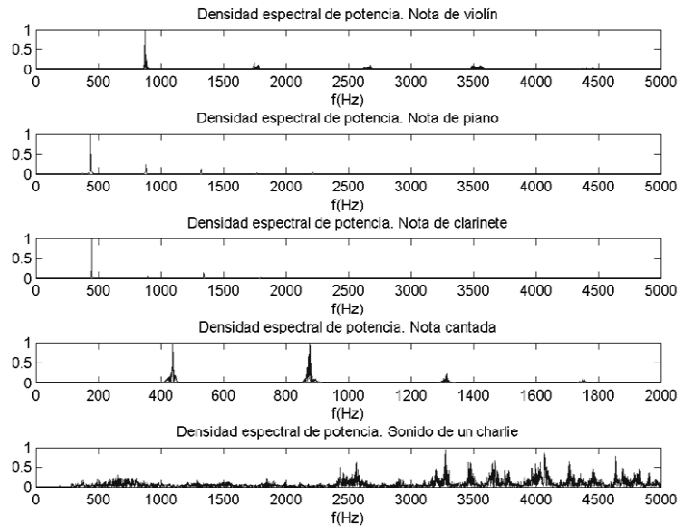
## Características temporales

- Instrumentos de entonación definida y no-definida



18

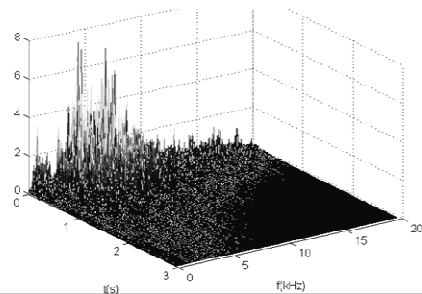
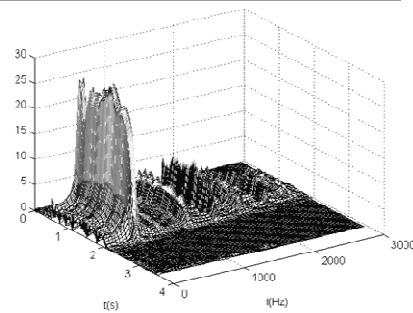
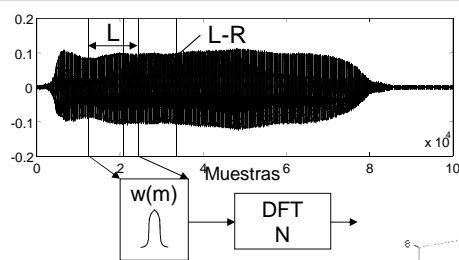
## Características frecuenciales



19

## Características frecuenciales

- Evolución temporal de las componentes
  - Transformada corta de Fourier (STFT)



20

## ***Organización de la charla***

---

1. Introducción
2. Técnicas de síntesis de sonido
3. Efectos de sonido
4. Conclusiones

21

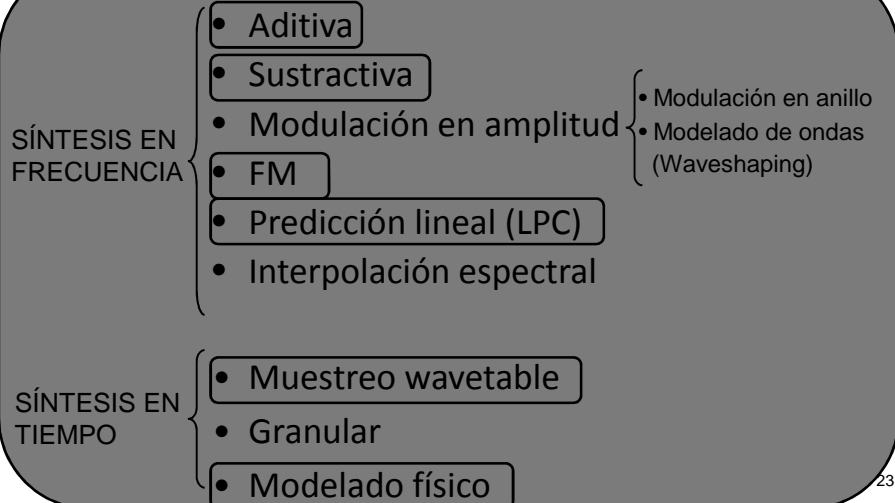
## ***Técnicas de síntesis de sonido***

---

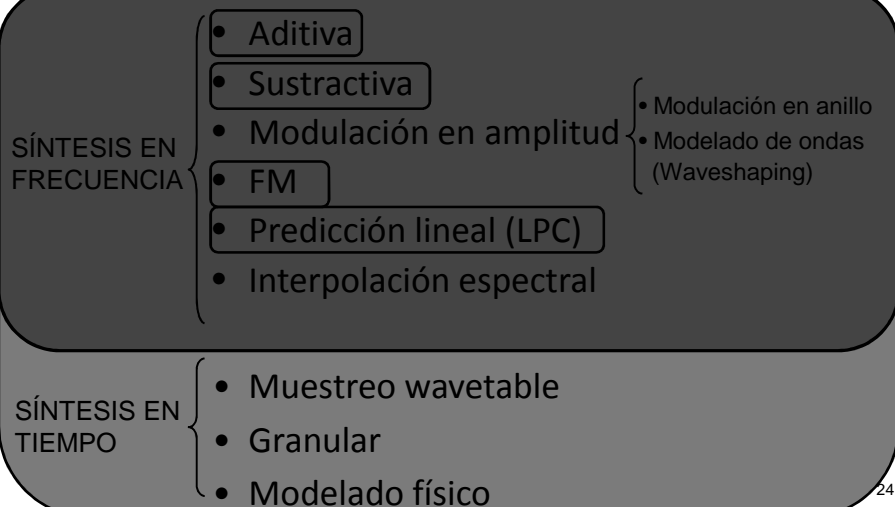
- Clasificación
- Síntesis en frecuencia
  - Aditiva
  - Sustractiva
  - FM
  - LPC
- Síntesis en tiempo
  - Wavetable
  - Modelado físico

22

## ***Técnicas de síntesis de sonido***



## ***Técnicas de síntesis de sonido***



## ***Síntesis en frecuencia: Aditiva***

- Fundamento: desarrollo en serie de Fourier de señales periódicas

$$x(t) = \sum_{k=0}^{\infty} A_k \text{sen}(k2\pi f_0 t + \phi_k)$$

donde:

$A_k$  es la amplitud de la componente 'k'

$f_0$  es la frecuencia fundamental de la señal

$\phi_k$  es la fase de la componente 'k'

- Problema:
  - Las señales musicales no son periódicas
  - No se puede sumar un número infinito de señales

25

## ***Síntesis en frecuencia: Aditiva***

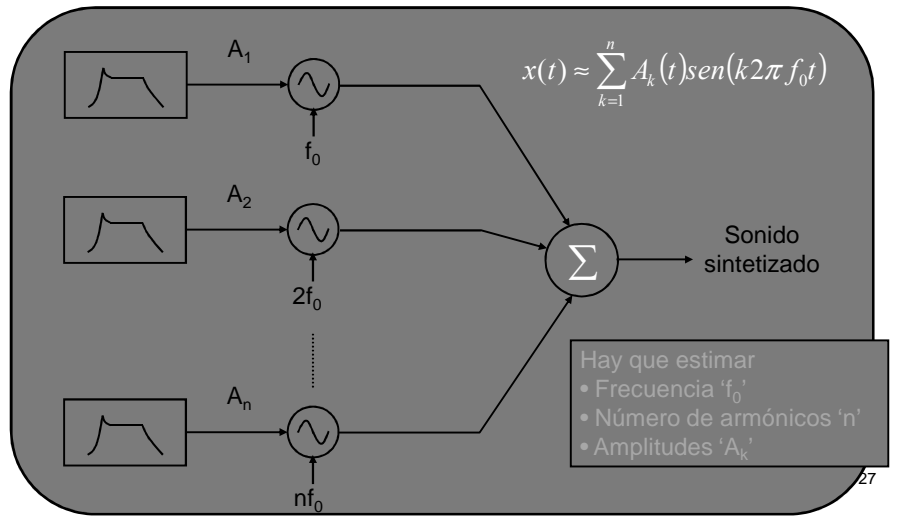
- Solución a la no periodicidad:
  - ↪ Aplicar STFT, de manera que se obtiene la evolución temporal de las componentes del espectro.
- Solución a la suma de un número infinito de componentes:
  - ↪ Eliminar las que están fuera del rango audible (20-20.000Hz).
- Simplificación:
  - ↪ Insensibilidad del oído a la fase de las componentes
- Expresión a utilizar:

$$x(t) \approx \sum_{k=1}^n A_k(t) \text{sen}(k2\pi f_0 t)$$

26

## Síntesis en frecuencia: Aditiva

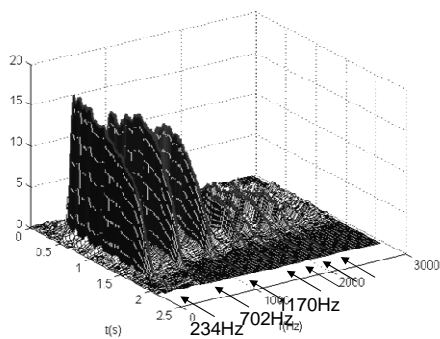
- Diagrama de bloques



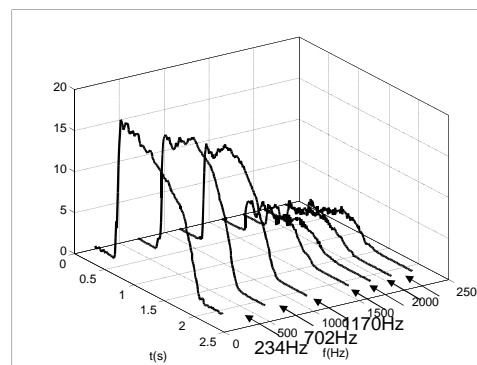
## Síntesis en frecuencia: Aditiva

- Ejemplo: Do4 de un clarinete

STFT del Do4 del clarinete



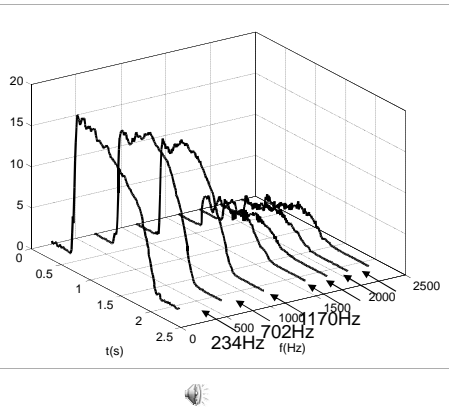
Aproximación STFT del Do4 del clarinete



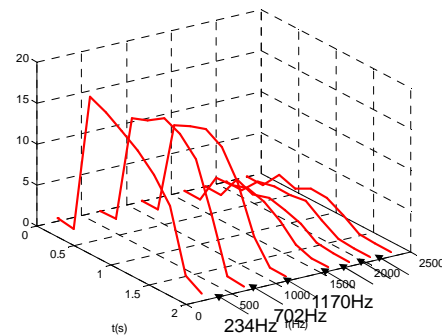
## Síntesis en frecuencia: Aditiva

- Ejemplo: Do4 de un clarinete

Aproximación STFT del Do4 del clarinete



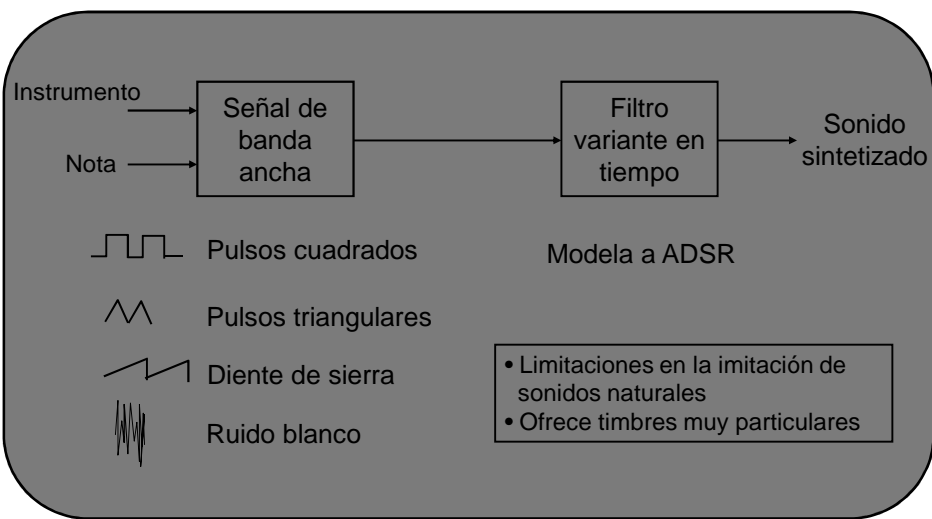
Aproximación de las envolventes



Aprox. Lineal Do<sub>4</sub> de un clarinete

## Síntesis en frecuencia: Sustractiva

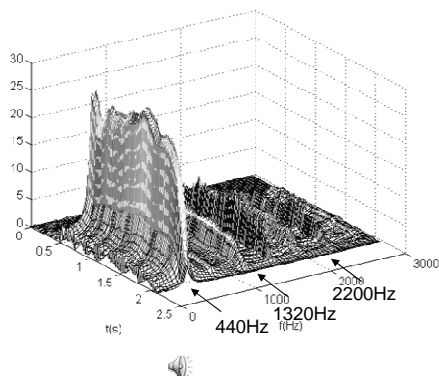
- Diagrama de bloques



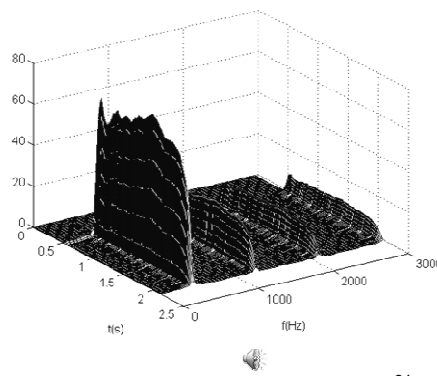
## Síntesis en frecuencia: Sustractiva

- La4 de un clarinete

STFT del La<sub>4</sub> de un clarinete  
(armónicos impares de 440Hz)



STFT de señal cuadrada  
(f=440Hz)



31

## Síntesis en frecuencia: FM

- Idea básica:
  - El timbre de una forma de onda simple, varía si se modula en frecuencia al dar lugar a una forma de onda más compleja.

- Señal modulada FM:

$$x_{FM}(t) = A_p \cos[\theta(t)] = A_p \cos[\omega_p t + \phi(t)] = A_p \cos\left[\omega_p t + 2\pi k \int A_m x_m(t) dt\right]$$

donde:

$$x_p(t) = A_p \cos(\omega_p t)$$

señal portadora

$$A_m x_m(t)$$

señal moduladora con  $x_m(t) \leq 1$

$$f_i(t) = \frac{1}{2\pi} \cdot \frac{d\theta(t)}{dt} = f_p + k A_m x_m(t)$$

frecuencia instantánea

$$k$$

constante de desviación en frecuencia

32



## Síntesis en frecuencia: FM

- Si la moduladora es un tono  $A_m \cos(\omega_m t)$

$$x_{FM}(t) = A_p \cos\left[\omega_p t + 2\pi k \int A_m \cos(\omega_m t) dt\right] = A_p \cos\left[\omega_p t + \frac{kA_m}{f_m} \sin(\omega_m t)\right]$$

$$I = \frac{kA_m}{f_m} = \frac{\Delta f}{f_m}$$

índice de modulación

$$\Delta f = kA_m$$

máxima desviación de frecuencia

$$f_i(t) = f_p + kA_m \cos(\omega_m t)$$

frecuencia instantánea

33

## Síntesis en frecuencia: FM

- La relación portadora-moduladora ( $f_p/f_m$ ), fija la posición de las componentes en el espectro



Se pueden tener espectros armónicos o inarmónicos

- El índice de modulación (I), determina la amplitud relativa de las componentes del espectro



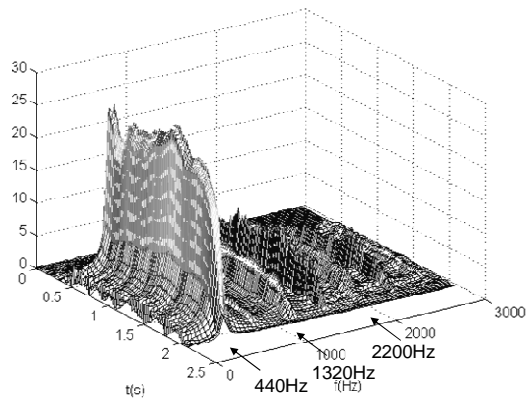
Su variación en el tiempo, hace que cambie la amplitud de las componentes así como su número

- Hay añadir la evolución temporal global de las notas: la envolvente espectral (ADSR)

34

## Síntesis en frecuencia: FM

- La4 de un clarinete



- El espectro es armónico y sólo están presentes los armónicos impares
- Los armónicos altos tienen poca energía

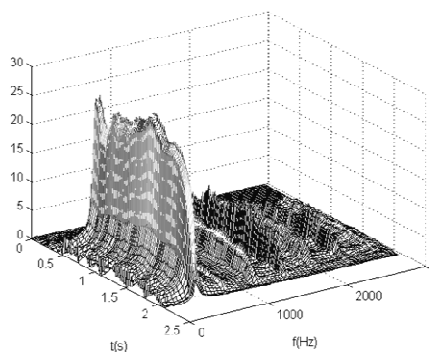


- Parámetros a definir:
  - $f_p/f_m$
  - Índice de modulación
  - Envolvente espectral

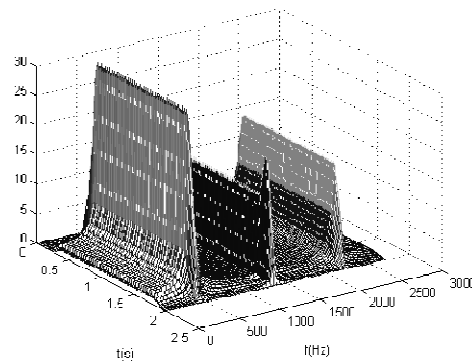
35

## Síntesis en frecuencia: FM

- La4 de un clarinete
- STFT del La<sub>4</sub> de un clarinete



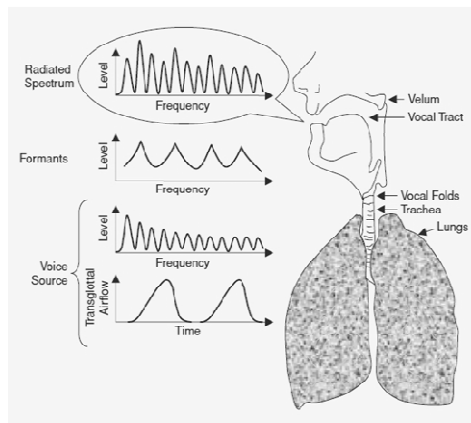
- STFT del La<sub>4</sub> de un clarinete sintetizado FM



36

## Síntesis en frecuencia: LPC

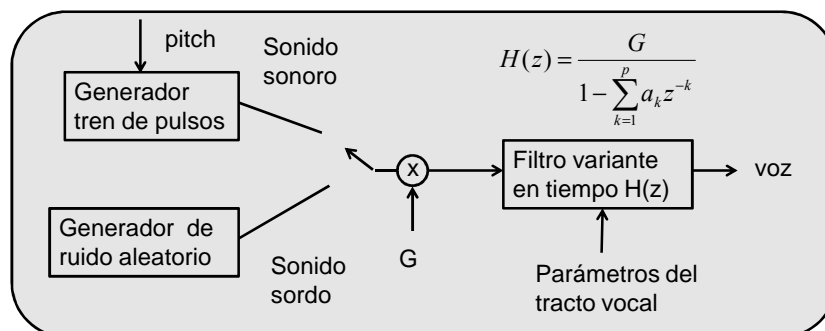
- LPC – Linear Predictive Coding. Muy utilizada en síntesis y reconocimiento de voz.
- Requerimientos básicos:
  - Establecer un modelo que represente los parámetros de la señal que se desea extraer.
  - Algoritmo de análisis basado en el modelo.
  - Programa de síntesis.



37

## Síntesis en frecuencia: LPC

- Modelo para el mecanismo de producción del habla



- Sonidos sonoros: Generados por las cuerdas vocales. Modificado por el tracto vocal. Espectro discreto y armónico.
- Sonidos sordos: producido por el aire a través de los estrechamientos del tracto vocal. Su espectro es continuo y de banda ancha.

38

## **Síntesis en frecuencia: LPC**

- Relación entre las muestras de voz y la excitación.

$$s(n) = \sum_{k=1}^p a_k s(n-k) + Gu(n)$$

$s(n)$  → muestras de la señal de voz

- Sistema de predicción lineal:

$$\tilde{s}(n) = \sum_{k=1}^p \alpha_k s(n-k)$$

- La función de un sistema de predicción lineal de orden  $p$  es:

$$P(z) = \sum_{k=1}^p \alpha_k z^{-k}$$

39

## **Síntesis en frecuencia: LPC**

- Error de predicción:

$$e(n) = s(n) - \tilde{s}(n) = s(n) - \sum_{k=1}^p \alpha_k s(n-k)$$

- El error de predicción es la salida de un sistema con función de transferencia:

$$A(z) = 1 - \sum_{k=1}^p \alpha_k z^{-k}$$

- Comparando el error  $e(n)$  con  $s(n) = \sum_{k=1}^p a_k s(n-k) + Gu(n)$



El filtro de predicción  $A(z)$  es el filtro inverso de  $H(z)$

40

## Síntesis en frecuencia: LPC

- Problema a resolver: determinar los coeficientes  $\{a_k\}$  de  $A(z)$

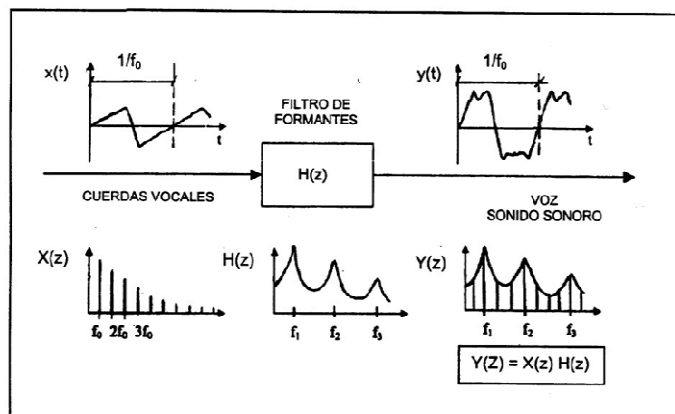
$$H(z) = \frac{G}{A(z)}$$

- La señal de voz es variante en tiempo  $\rightarrow$  los coeficientes deben estimarse para pequeños segmentos de la señal de voz.
- Encontrar  $\{a_k\}$  que minimice el error cuadrático medio de predicción en un pequeño fragmento de la forma de onda de voz.
- Determinado  $\{a_k\} \rightarrow$  calcular el error  $\rightarrow$  utilizarlo como entrada del filtro  $H(z) = G / A(z) \rightarrow$  reconstruir la señal de voz  $s[n]$ .
- Si la predicción lineal está funcionando bien  $\rightarrow$  error parecido a un tren de impulsos (un impulso por trama).
  - Señal fuente  $\rightarrow$  un tren de impulsos ideal para sonoros
  - Ruido  $\rightarrow$  para fricativos.

41

## Síntesis en frecuencia: LPC

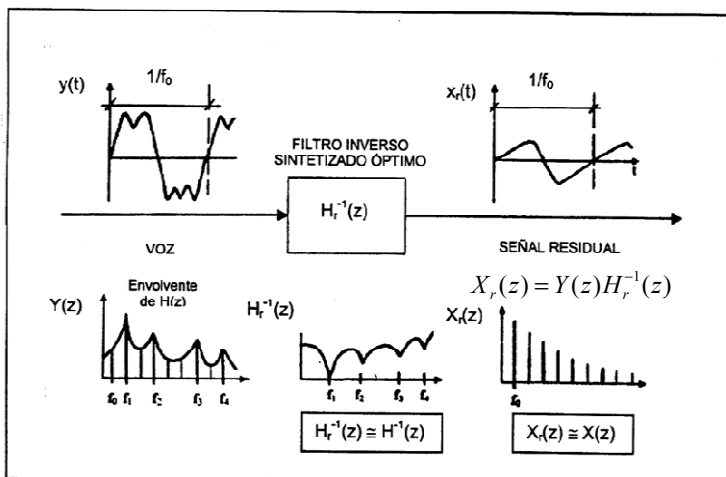
- Emisión de sonidos sonoros.



42

## Síntesis en frecuencia: LPC

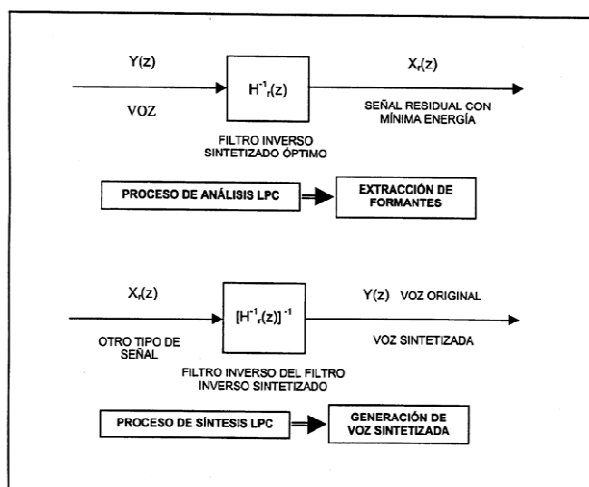
- Emisión de sonidos sonoros.



43

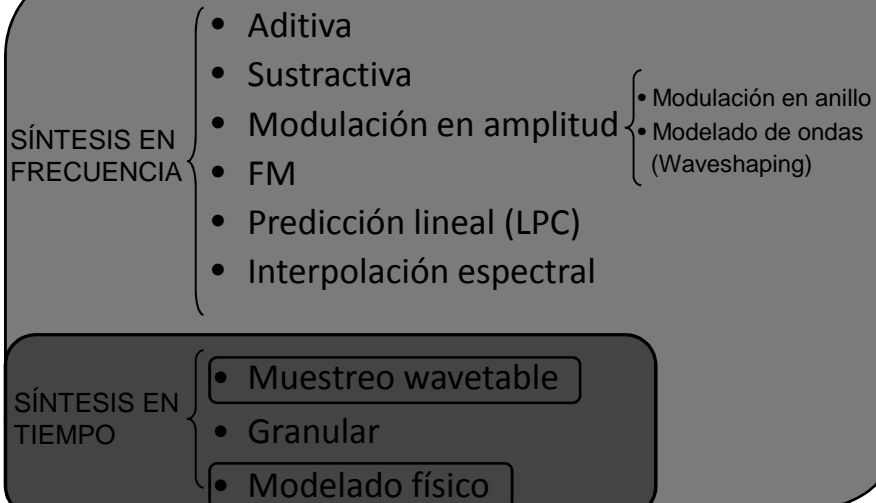
## Síntesis en frecuencia: LPC

- Emisión de sonidos sonoros.



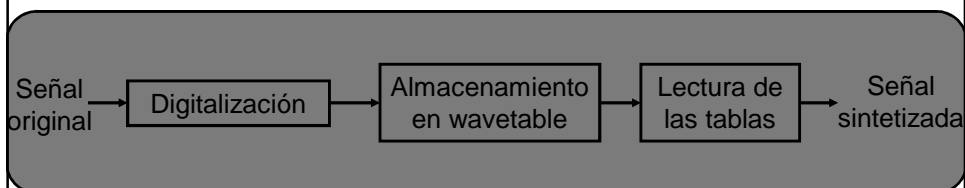
44

## ***Técnicas de síntesis de sonido***



## ***Síntesis en tiempo: Wavetable***

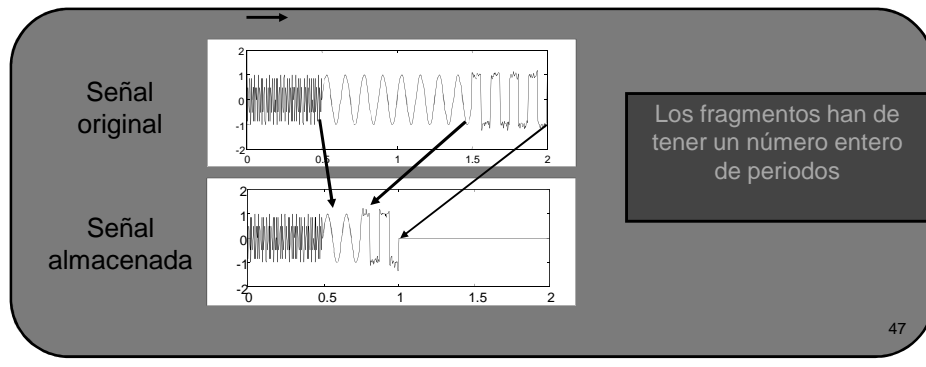
- Fundamento: Digitalizar el sonido original y almacenar las muestras en una tabla de onda (Wavetable)



Necesidad de optimizar la cantidad de información a almacenar

## ***Síntesis en tiempo: Wavetable***

- Ejemplo de almacenamiento para instrumentos de entonación definida.
  - Fragmentos de señal que se repiten periódicamente (sostenimiento) → sólo se almacena un periodo
  - Fragmentos de señal no repetitivos → se almacenan completos



## ***Síntesis en tiempo: Wavetable***

- Ejemplo para un piano. Se almacenan dos partes del sonido del instrumento:
  - Se almacena el ataque → sonido del martillo golpeando una cuerda del piano
  - Se almacena un pequeño fragmento del periodo de sustain

Cuando se toca la nota:

- La parte del ataque se reproduce, seguido por una repetición en bucle del segmento de sostenimiento.
- El segmento que se repite, se procesa para crear un decaimiento natural del sonido

48



## ***Síntesis en tiempo: Wavetable***



---

- Ventajas:
  - Fácil de implementar
  - Muy buena calidad
  - Eficiente en el tratamiento de señales periódicas
  
- Inconvenientes:
  - No es muy útil para crear sonidos nuevos
  - Necesita mucha memoria para almacenar sonidos muestreados

49

## ***Síntesis en tiempo: Modelado físico***

---

- Síntesis de sonido utilizando un modelo matemático para simular las propiedades físicas de la fuente del sonido.
  
- Cook la utilizó para sintetizar voz.
  
- El sonido se genera usando:
  - Resonadores: cómo vibran los elementos físicos del instrumento
    -  parámetros que describen los materiales físicos empleados en el instrumento
  - Generadores: excitación que produce el sonido
    -  interacción del usuario con el instrumento.
  
- Método muy utilizado: Diseño de guías de onda a partir de modelos discretos en tiempo.

50

## ***Calidad de los sonidos sintetizados***

- Proceso subjetivo de evaluación

### **Categorías de calidad de sonido**

Mejor ↓ Peor	• Físicamente indistinguible	→	Misma forma de onda
	• Percepción indistinguible	→	Diferencia inaudible
	• Musicalmente indistinguible	→	Sonidos intercambiables
	• Musicalmente aceptable	→	Cambio de interprete
	• Musicalmente diferente	→	Instrumento distinto
	• Musicalmente independiente	→	Sonidos diferentes

51

## ***Organización de la charla***

1. Introducción
2. Técnicas de síntesis de sonido
3. Efectos de sonido.
4. Conclusiones

52

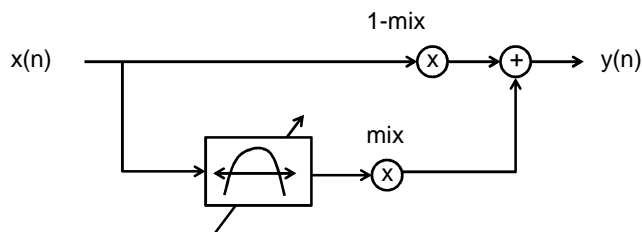
## **Digital Audio Effects**

- Efectos de audio. Utilizados por individuos involucrados en la generación de señales musicales.
  - Formas especiales de tocar de los músicos
  - Usos concretos de los micrófonos
  - Variaciones múltiples de la señal de audio
  
- Clasificación:
  - Filtrado — Paso alto, paso bajo, ecualizadores
  - Filtros variantes en tiempo — Wah-wah, Phaser
  - Efectos de retardo — Vibrato, Flanger, Chorus, Echo
  - Moduladores — Ring modulation, Tremolo, Vibrato
  - Procesado no-lineal — Compresión, Limitadores, Distorsión
  - Efectos espaciales — Panning, Reverb, Surround Sound

53

## **Filtro variante en t: wah-wah**

- Filtros variantes en tiempo
  - Wah-wah. Filtro paso banda con frecuencia central variable y ancho de banda estrecho. La señal filtrada se mezcla con la señal original.

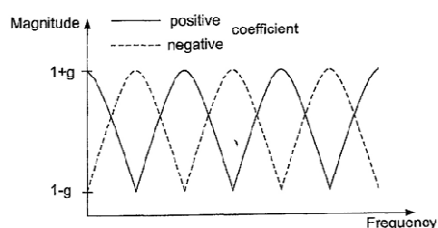
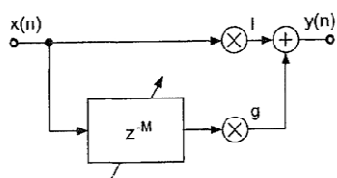


54

## Efectos de retardo

- Efectos de retardo — Vibrato, Flanger, Chorus, Echo
- Estructuras básicas con filtros FIR e IIR muy sencillos.
- Universal Comb Filter – combinación de FIR e IIR.
- FIR Comb Filter

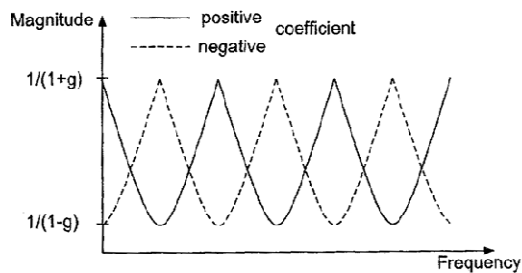
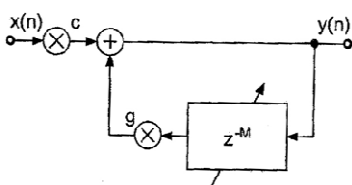
$$y(n) = x(n) + gx(n - M) \quad \text{con} \quad M = \tau / f_s$$



## Efectos de retardo

- IIR Comb Filter  
Simula una reflexión sin fin.

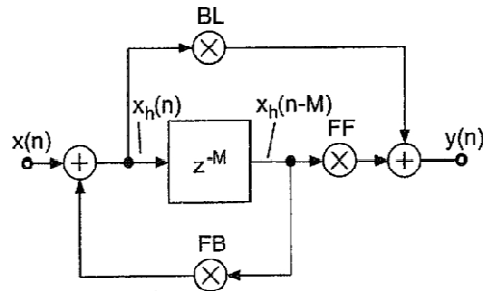
$$y(n) = Cx(n) + gy(n - M) \quad \text{con} \quad M = \tau / f_s$$



56

## Efectos de retardo

- Universal Comb Filter

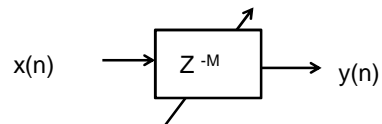


	BL	FB	FF
FIR comb	1	0	g
IIR comb	1	g	0
allpass	a	-a	1
delay	0	0	1

57

## Efectos de retardo: Vibrato

- Variación periódica del retardo temporal
- Implementación: un retardo y un oscilador de baja frecuencia para variar el retardo.

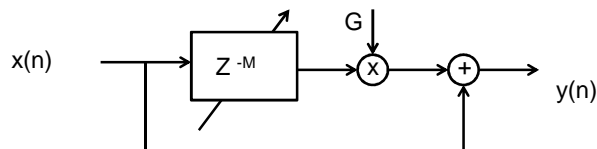


- Retardo típico: 5 a 10 ms
- Oscilador de baja frecuencia: 5-14Hz

58

### ***Efectos de retardo: Flanging***

- Variación periódica del retardo temporal donde se suma también la señal original.
- Implementación

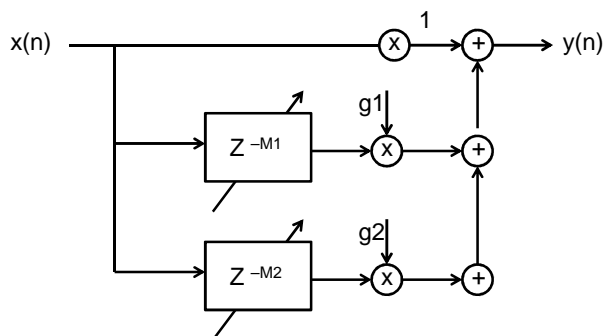


- Retardo típico: >15 ms
- Oscilador de baja frecuencia: 1Hz

59

### ***Efectos de retardo: chorus***

- Simula varias voces cantando en un unísono imperfecto.
- Implementación: Se suman varias copias de la señal de entrada retrasadas y con variaciones aleatorias de dicho retardo.



- Retardo típico: 10-25 ms.
- Retardo varía aleatoriamente poco y despacio.

60

## ***Efectos de retardo: tabla resumen***

- Configuración de varios efectos de retardo

Efecto	Rango de retardos (ms)	Modulación
Resonador	0...20	Ninguna
Flanger	0...15	Sinusoidal
Chorus	10...25	Aleatoria
Slapback	25...50	Ninguna
Echo	>50	Ninguna

61

## ***Efectos moduladores***

- Efectos de modulación:

Una señal de audio modula en amplitud y fase una señal sinusoidal

- Efecto Wah-wah → modulación de amplitud
- Efectos Vibrato, Chorus y Flanger → modulación de fase


- Ring Modulation:  $y(n) = x(n) \cdot m(n)$

 Voz normal     Voz robot     Audio ring\_mod

- Modulación de amplitud: Vibrato

$$y(n) = (1 + \alpha m(n)) \cdot x(n)$$

 Tremolo con seno

 Tremolo con onda triangular <sub>62</sub>

## ***Efectos espaciales***

---

- Efectos de espaciales: Cambian la localización espacial del sonido.
- Ejemplo: Introducción de la respuesta al impulso de un determinado espacio.
  - Grabar la respuesta a un pequeño impulso en el recinto (una palmada, un golpe de tambor).
  - Convolucionar el sonido con la respuesta al impulso del recinto.

63

## ***Conclusiones***

---

- Cómo caracterizar una nota musical.
- Distintas técnicas de síntesis de sonido:
  - Frecuencia
  - Tiempo
- Diversas formas de introducir efectos a una señal musical.
  - Filtrado
  - Modulación
  - Efectos espaciales

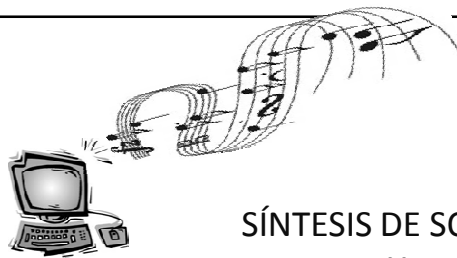
64



## **Bibliografía**

- M. Russ, "Sound Synthesis and Sampling", Third edition, Focal Press, 2009.
- U. Zölzer, "DAFX. Digital Audio Effects", Wiley, 2011.
- R. Boulanger, "The Csound Book", MIT Press, 2001.
- D. Simón Zorita, "Análisis y Síntesis de Señales Musicales", E.U.P. Madrid, 2000.
- U. Zölzer, "Digital Audio Signal Processing", Second edition, John Wiley, 2008.
- S. Bilbao, "Numerical Sound Synthesis", John Wiley, 2009.
- L. R. Rabiner, R. W. Schafer, "Digital Processing of Speech Signals", Prentice-Hall, 1978.
- P. R. Cook, "Singing voice synthesis: History, current work, and future directions," Computer Music Journal, vol. 20, pp. 38-46, Fall 1996.
- [https://ccrma.stanford.edu/~jos/pasp/Digital\\_Waveguide\\_Model.html](https://ccrma.stanford.edu/~jos/pasp/Digital_Waveguide_Model.html)
- Bonada, J., Serra, X., "Synthesis of the Singing Voice by Performance Sampling and Spectral Models", IEEE Signal Processing Magazine, vol.24, no. 2, pp. 67-79, 2007.

65



## SÍNTESIS DE SONIDO Y EFECTOS DE AUDIO DIGITAL

Isabel Barbancho  
Profesora Titular de Universidad  
Universidad de Málaga  
E.T.S. Ingeniería de Telecomunicación  
Campus de Teatinos s/n, 29071 Málaga  
Teléfono: 952132587  
email: [ibp@ic.uma.es](mailto:ibp@ic.uma.es)

Noviembre, 2011

66