

Joaquín Luque Rodríguez

**FUNDAMENTOS FÍSICOS
DEL HABLA**
Segunda edición

**UNIVERSIDAD DE SEVILLA
DEPARTAMENTO DE TECNOLOGÍA ELECTRÓNICA**

Joaquín Luque Rodríguez

**FUNDAMENTOS FÍSICOS
DEL HABLA**
Segunda Edición

Universidad de Sevilla
Departamento de Tecnología Electrónica
Servicio de Publicaciones
Sevilla, 1994
Segunda edición, 2002

Escuela Técnica Superior de Ingeniería Informática
Avenida Reina Mercedes s/n
41012-Sevilla. SPAIN.
(+34) 95.455. 27.86
jluque@us.es

ÍNDICE

1	INTRODUCCIÓN	2
2	CARACTERIZACIÓN TEMPORAL DEL SONIDO	2
2.1	MARCO CONCEPTUAL.....	2
2.2	SOPORTE MATERIAL DEL SONIDO.....	3
2.3	ANÁLISIS TEMPORAL DE UNA VIBRACIÓN SIMPLE.....	4
2.4	ANÁLISIS TEMPORAL DE UNA VIBRACIÓN COMPLEJA.....	8
2.5	EL SONIDO COMO VIBRACIÓN TEMPORAL COMPLEJA.....	12
3	CARACTERIZACIÓN ESPECTRAL DEL SONIDO	13
3.1	ESPECTRO FRENTE A TIEMPO.....	13
3.2	ANÁLISIS ESPECTRAL DE UNA VIBRACIÓN SIMPLE.....	14
3.3	ANÁLISIS ESPECTRAL DE UNA VIBRACIÓN COMPLEJA.....	15
3.4	EL SONIDO COMO VIBRACIÓN ESPECTRAL COMPLEJA.....	18
3.5	ESPECTROS VARIABLES.....	19
4	BIBLIOGRAFÍA	21

1 INTRODUCCIÓN

La estructura del lenguaje, su clasificación e interpretación ha sido una constante inquietud del pensamiento occidental. ¿Cómo es posible el lenguaje? ¿Qué rasgos caracterizan el lenguaje humano frente al animal? ¿Cuáles son las características comunes de la comunicación humana en todas las lenguas? La respuesta a estas preguntas ha ocupado, y sigue ocupando, a buen número de lingüistas, filósofos, psicólogos, científicos y, en general, pensadores de las más diversas áreas del saber.

A este interés general por el lenguaje ha venido a sumarse en las últimas épocas, con interés creciente, el problema del lenguaje oral. ¿Qué características diferenciales posee el lenguaje oral humano? ¿Cuál es su relación con el lenguaje escrito? Y de manera más específica aún, con el auge de los computadores, ¿cuáles son los parámetros esenciales de un sonido que permiten su interpretación como un determinado mensaje oral? ¿Cómo puede reconocerse el habla? Y, también en sentido inverso, ¿cómo se debe generar un sonido para que un hombre lo interprete como un mensaje oral normal? Son en suma, los problemas del reconocimiento y síntesis del habla, entendida ésta como lenguaje oral.

Casi cualquier intento de aproximación a los problemas planteados debe pagar el tributo de realizar un estudio previo de los principios físicos que subyacen al fenómeno del habla. Han de conocerse las características que, a la luz de la ciencia contemporánea, presenta cualquier sonido en general y el habla en particular. Debe estudiarse también el proceso de emisión (habla) y recepción (escucha) del lenguaje oral, y deben clasificarse las características **físicas** de estos sonidos que permitan su reconocimiento y su síntesis.

En este documento se encara la cuestión de presentar los principios físicos que dan soporte al fenómeno del lenguaje oral, tratando de guardar el equilibrio entre dos enfoques pretendidamente complementarios y no pocas veces contrapuestos: por un lado un enfoque humanístico en el que se abordan estas cuestiones muy superficialmente, y a veces con escaso rigor, profundizando en clasificaciones, descripciones e interpretaciones; y por otro lado un enfoque tecnicista en el que la complejidad de la formulación físico-matemática hace perder, en no pocas ocasiones, la perspectiva de conjunto y que, en el mejor de los casos, impide el acceso a los no iniciados.

2 CARACTERIZACIÓN TEMPORAL DEL SONIDO

2.1 Marco conceptual

Nos encontramos inmersos en una realidad sonora. Percibimos el sonido como una realidad incontestable, como algo ajeno a nosotros, como algo otro, como algo real. Pero, ¿cuál es la íntima realidad del sonido? ¿En qué consiste **esencialmente** un sonido? La respuesta a esta pregunta (como casi a la de cualquier otra pregunta) ha sido diversa a lo largo de la historia. Pero incluso hoy día podría enfocarse desde distintos ángulos dependiendo de la formación o intereses desde la que se responda. En mi opinión, y en la de la mayoría de los autores que abordan la cuestión (incluso los de tradición humanista), el enfoque más adecuado para describir un sonido en el ámbito que nos ocupa es, sin duda, el enfoque físico-matemático.

Pero, ¿en qué consiste este enfoque? O más precisamente, ¿cuál es su verdad? En mi opinión, la ciencia nos provee con una serie de conceptos que están basados en una aprehensión primordial de la realidad pero que la trascienden, acercándonos a lo que la realidad es *en realidad*. Partimos de una percepción simple, no problemática, de un sonido y teorizamos sobre lo que ese sonido sea *en realidad*. Ayudándonos de nuestra percepción y comprensión de otras realidades, construimos conceptos y teorías que nos hacen vislumbrar, sólo vislumbrar, la *esencia* del sonido. Ha habido otras construcciones teóricas en el pasado pero nuestra teoría actual nos ofrece más confianza. En el conjunto de cuestiones que conocemos y sabemos, es la explicación más consistente, más coherente, más útil. Nos permite describir, explicar y predecir. Nos permite modificar y construir. Esa es su verdad. A ella dedicaremos las próximas líneas.

2.2 Soporte material del sonido

Podemos decir que un sonido es una vibración audible. Para ello requerimos al menos tres elementos:

- a) Una fuente de producción del sonido que es, en todos los casos, un elemento material en vibración, es decir, en movimiento de vaivén.
- b) Uno o varios medios materiales que propaguen la vibración desde la fuente del sonido hasta el receptor.
- c) Un aparato auditivo (real o potencial) que sea capaz de recibir las vibraciones y transformarlas en sensaciones. En el hombre este aparato auditivo consta de múltiples órganos, desde el pabellón auditivo externo (la oreja) hasta el cerebro, o si se quiere, la mente.

Veamos con un ejemplo el aspecto material del sonido. Imaginemos una guitarra. Con el pulgar de la mano derecha tomamos una de sus cuerdas, digamos la superior, y la desplazamos ligeramente de su posición de reposo. A continuación soltamos. Podemos observar (visual y acústicamente) que la cuerda comienza a moverse arriba y abajo de su posición de equilibrio con un movimiento vibratorio. Ya tenemos la fuente del sonido: una cuerda (material) que vibra.

Pero una guitarra en el vacío no suena. Se necesita el aire, u otro material, que propague el sonido hasta el receptor. ¿Cómo se produce esta propagación? El mecanismo por el que esto se produce es relativamente simple. La parte externa de la cuerda, al desplazarse hacia arriba *empuja* a las *gotas* (moléculas) de aire que están en su proximidad. Éstas a su vez empujan a las que se encuentran a su lado, y éstas últimas a sus vecinas. Se va produciendo así un fenómeno parecido al denominado *efecto dominó*: al empujar la primera ficha de dominó de una hilera ésta va empujando a la siguiente hasta que el movimiento se transmite a la última ficha. De hecho las *gotas* de aire funcionan como una hilera de fichas de dominó que estuvieran unidas por pequeñas tiras de goma elástica. Esto hace que la hilera no sólo se desplace en un sentido, sino que pueda hacerlo también en sentido inverso. En efecto, cuando la cuerda de la guitarra, en su movimiento vibratorio, termina de desplazarse hacia arriba y comienza a desplazarse hacia abajo *tira* de las gotas de aire circundantes, éstas de sus vecinas próximas y éstas de las suyas. Igual que las fichas de dominó sujetas con la goma. De esa forma el movimiento vibratorio de la cuerda se

propaga a través del aire (o de cualquier otro medio material) hasta otro punto del espacio. Tenemos el segundo de los elementos del sonido: el medio de propagación.

Por último, en el aparato auditivo, las vibraciones de las gotas de aire externo se propagan, de manera similar a como se acaba de describir, y alcanzan los nervios auditivos que mediante complejas operaciones bioquímicas transportan al cerebro la información recibida. En éste último, mediante mecanismos que todavía hoy nos son desconocidos, se produce la audición, es decir, la sensación mental de sonido. Tenemos el tercero y último de los elementos enunciados más arriba.

2.3 Análisis temporal de una vibración simple

Ya hemos visto que el sonido, en su aspecto físico, es fundamentalmente un fenómeno vibratorio. Pero para poder distinguir, clasificar, caracterizar y, en definitiva, operar con vibraciones debemos describirlas con mayor precisión que lo hecho hasta ahora. Sabemos que el sonido de la guitarra, y por tanto la vibración en que se materializa, no es el mismo si pulsamos la sexta cuerda o la primera. Notamos una diferencia de tono. También somos capaces de distinguir la diferencia entre el sonido que emite una cuerda de nuestra guitarra al *acariciarla* y el que emite al pulsarla con más fuerza (al desplazarla más de su posición de equilibrio). Percibimos diferencias de intensidad. Estas son las dos características principales de un sonido y de una vibración: tono e intensidad. ¿En qué consisten? ¿Cómo podemos conceptualizar eso que ya hemos aprehendido como real?

Se pueden realizar múltiples conceptualizaciones de los fenómenos que percibimos como tono e intensidad o, más generalmente, caben varias formas de conceptualizar una vibración. En el campo de estudio que nos ocupa son dos las representaciones que se realizan de la misma realidad: a) una representación temporal; y b) una representación espectral. Veremos ahora con detenimiento la primera de ella y más adelante nos referiremos a la segunda.

La representación temporal o análisis temporal de una vibración es quizás la forma más inmediata y, si se quiere, la más intuitiva de estudiar una vibración. Como numerosos autores apuntan, sólo podemos percibir en el espacio y el tiempo. No podemos tener percepciones *a*-temporales y *a*-espaciales. Sin embargo sí son posibles conceptos de esta naturaleza. Por ejemplo los conceptos de amor, libertad, o número, no reclaman la espacialidad ni la temporalidad. Por tanto, la representación temporal de la vibración, si bien es la más cercana a la percepción (que siempre tiene carácter temporal) no será la única posible.

Para comenzar el análisis temporal de las vibraciones supongamos que tenemos un bola metálica sujeta mediante dos muelles iguales a un marco de madera que está fijo en el suelo (figura 1). Originalmente la bola se encuentra a 30 centímetros del suelo.

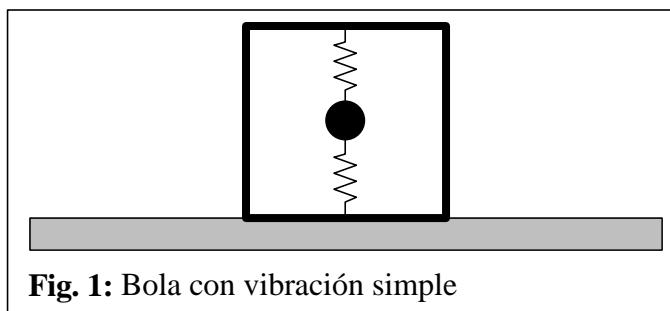
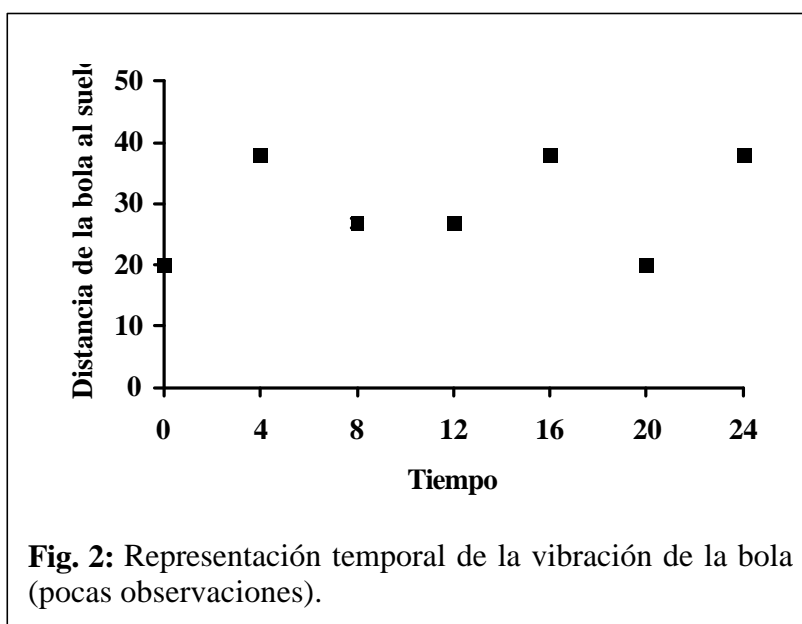
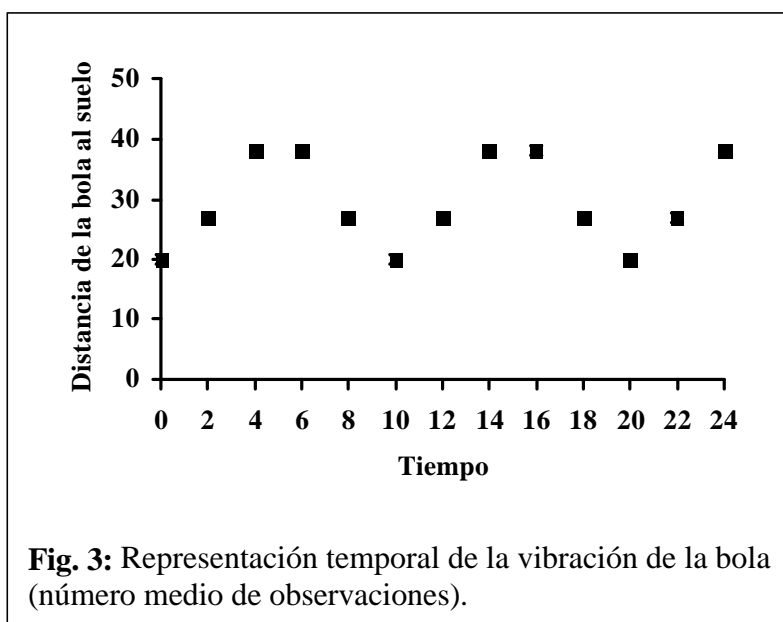


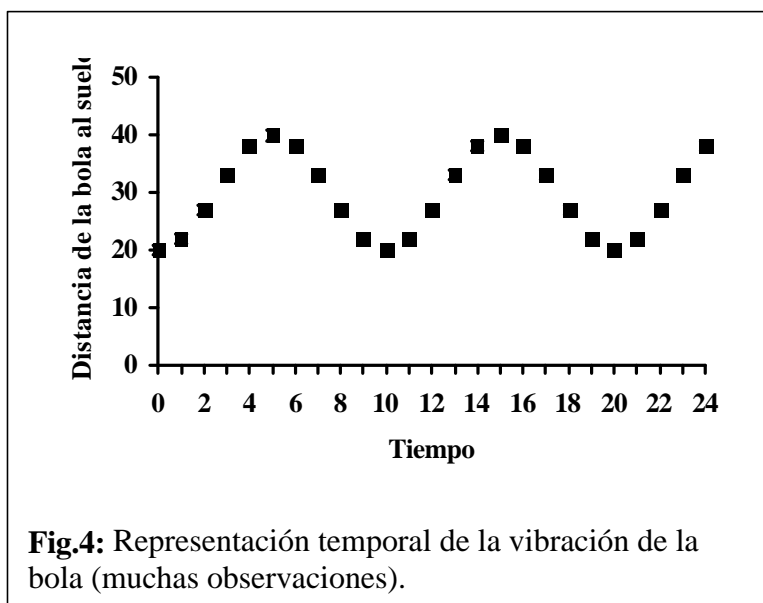
Fig. 1: Bola con vibración simple

Si desplazamos 10 centímetros la bola hacia abajo y la soltamos, comenzará a vibrar, arriba y abajo con un movimiento que se denomina movimiento armónico simple. Por las características físicas del artilugio, este movimiento vibratorio produce un sonido que percibimos como un zumbido bastante grave. Si nos fijamos en la posición que ocupa la bola (distancia de la bola al suelo) en diferentes instantes de *tiempo* estamos haciendo un análisis *temporal* del fenómeno vibratorio. Esto puede representarse de manera gráfica de acuerdo con la figura 2. En ese gráfico aparecen las posiciones de la bola en varios momentos.

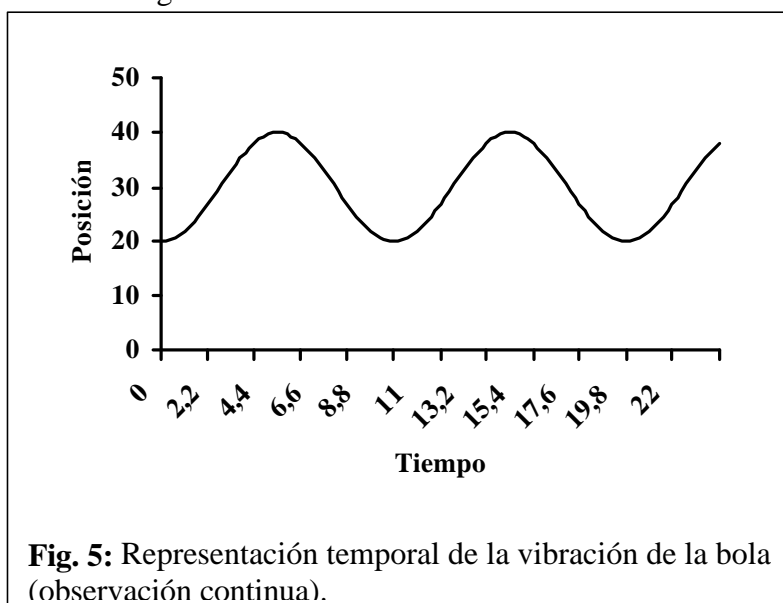


Si quisiésemos representar la posición de la bola en instantes intermedios tendríamos que fijarnos en (y dibujar) la posición de la bola en *más* instantes de tiempo, como por ejemplo en las figuras 3 y 4.

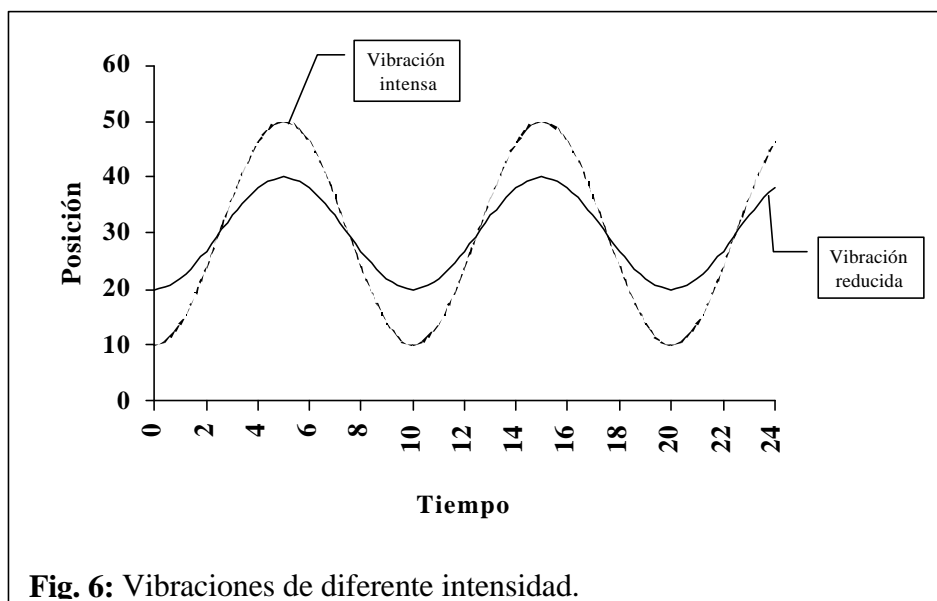




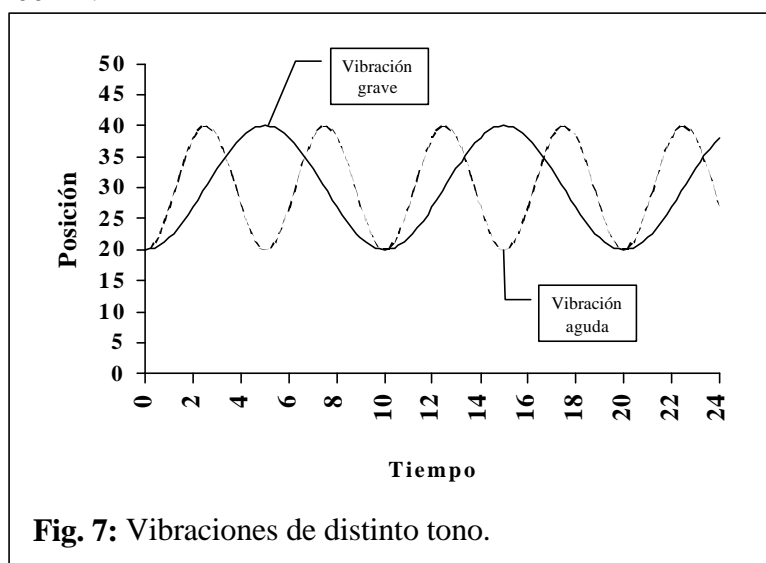
Por último si queremos saber la posición de la bola en *todo* momento nos encontramos con una gráfica como la de la figura 5.



En estas gráficas podemos ya identificar los aspectos de intensidad y tono de la vibración. En efecto, se denomina intensidad (técnicamente amplitud) a la máxima desviación de la bola de su posición de reposo. En este caso la amplitud o intensidad de la vibración es de 10 centímetros. Si ahora tiramos de la bola hacia abajo 20 centímetros antes de soltarla, percibimos una vibración más intensa que, de acuerdo con nuestra conceptualización, se puede medir por una amplitud de 20 centímetros (figura 6). Pero, ¿qué es el tono?



El movimiento vibratorio de la bola es periódico, es decir, la posición de la bola se va repitiendo una y otra vez a intervalos regulares. Este hecho queda claramente de manifiesto en las gráficas anteriores en las que vemos cómo la bola alcanza siempre la misma posición, por ejemplo el punto más alto, a intervalos constantes de 10 milisegundos (milésimas de segundo), o lo que es lo mismo, cada centésima de segundo. A esta duración de la espera entre dos repeticiones de la misma posición de la bola se la denomina periodo de vibración y en este caso vale 10 milisegundos. La magnitud inversa del periodo se denomina *frecuencia*. Si la bola realiza un movimiento de vaivén cada décima de segundo, ¿cuántos de dichos movimientos realiza en un segundo? Obviamente realiza 100 vibraciones en un segundo y por tanto, su frecuencia es de 100 vibraciones por segundo, 100 ciclos por segundo o, como se denomina técnicamente, 100 *Hercios* (abreviadamente se escribe 100 Hz.). Pues bien el tono de la vibración (el tono del zumbido percibido) es, precisamente, su frecuencia: 100 Hz.



Si sustituimos los muelles que sujetaban a la bola por otros más rígidos (más duros),

sabemos que al desplazar la bola 10 centímetros y soltarla va a oscilar más rápidamente, con mayor *frecuencia*. Digamos que su frecuencia será de 200 Hz., es decir, que realizará 200 vibraciones por segundo. Esto puede verse gráficamente en la figura 7. El tono que ahora percibimos es más agudo que en el caso anterior: un zumbido de 200 Hz.

2.4 Análisis temporal de una vibración compleja

En el apartado anterior hemos expuesto las dos características fundamentales de un sonido (o de una vibración): su intensidad y su frecuencia. Pero, el caso que estudiamos era el más simple posible. No todas las vibraciones (y de ellas casi ningún sonido) pueden representarse por un modelo tan simplificado. Si bien es verdad que intensidad y tono son los principales rasgos de un sonido, no son los únicos pues, ¿en qué se diferencia un *La* de guitarra del mismo *La* de piano? Podemos hacer que su intensidad y su tono sean iguales pero sin embargo son perfectamente distinguibles. La característica que los diferencia se denomina *timbre*. Para poder conceptualizar adecuadamente el fenómeno del timbre debemos abandonar el modelo simplificado de vibración usado hasta el momento y adentrarnos en lo que se denomina vibración o sonido complejo.

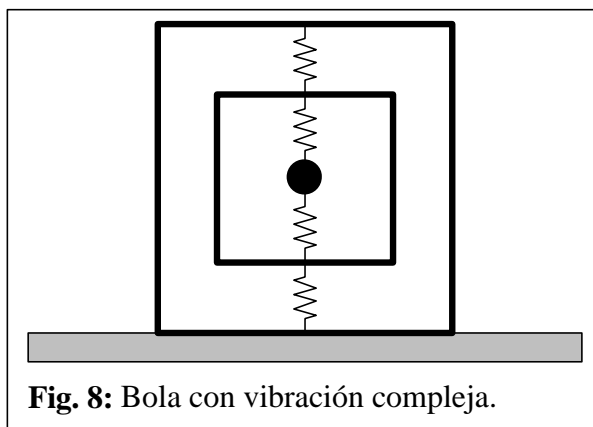


Fig. 8: Bola con vibración compleja.

Consideremos de nuevo la bola vibrante de nuestro ejemplo anterior. Supongamos ahora que el marco al que se sujeta la bola no está fijo en suelo sino que, a su vez, se halla sujeto a otro marco exterior por un par de muelles. El segundo marco se encuentra anclado al suelo (figura 8). Con esta configuración tenemos dos posibles movimientos vibratorios: el de la bola respecto al marco interior; y el del marco interior con respecto al suelo. El movimiento de la bola con respecto al suelo será pues la combinación de dos vibraciones. Si sujetamos el marco interior y desplazamos y soltamos la bola, su movimiento con respecto al suelo será, como en el caso anterior, un movimiento vibratorio simple (figura 9).

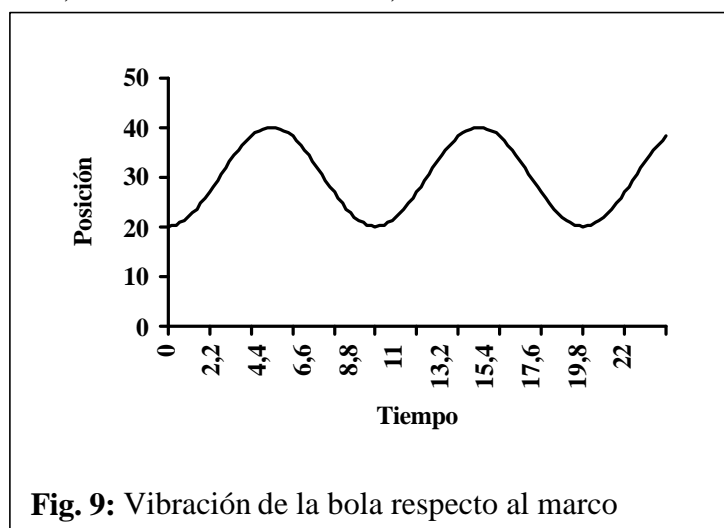
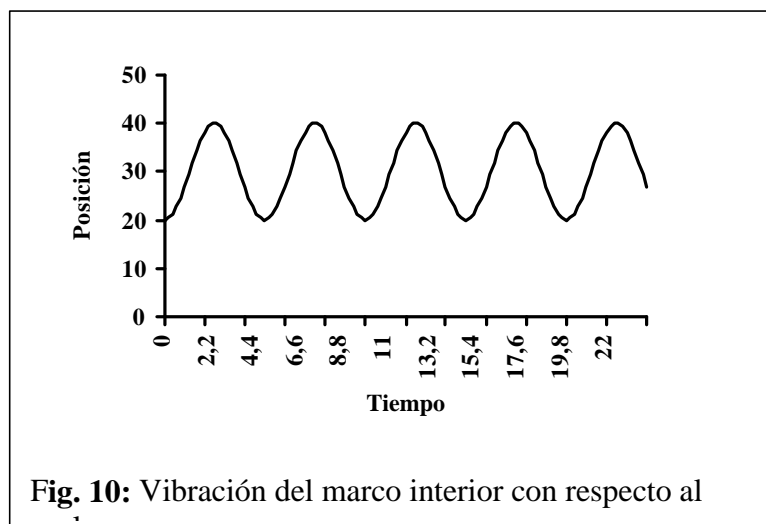
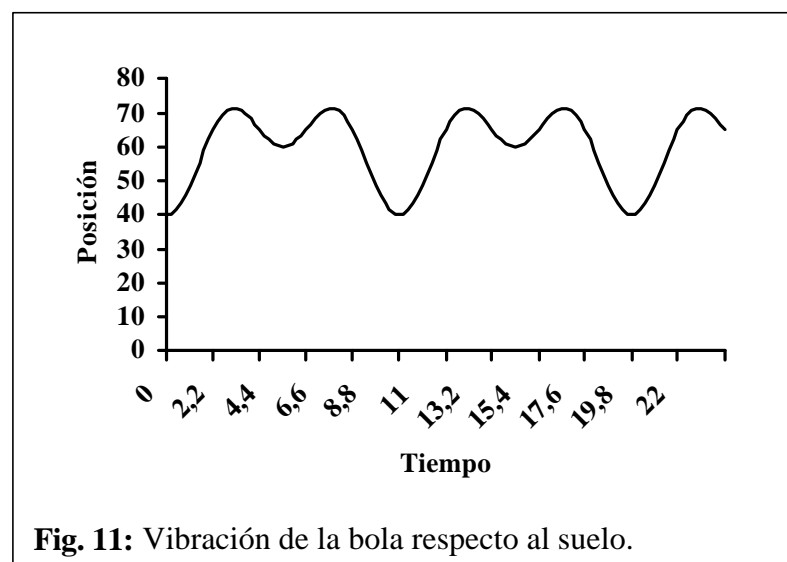


Fig. 9: Vibración de la bola respecto al marco

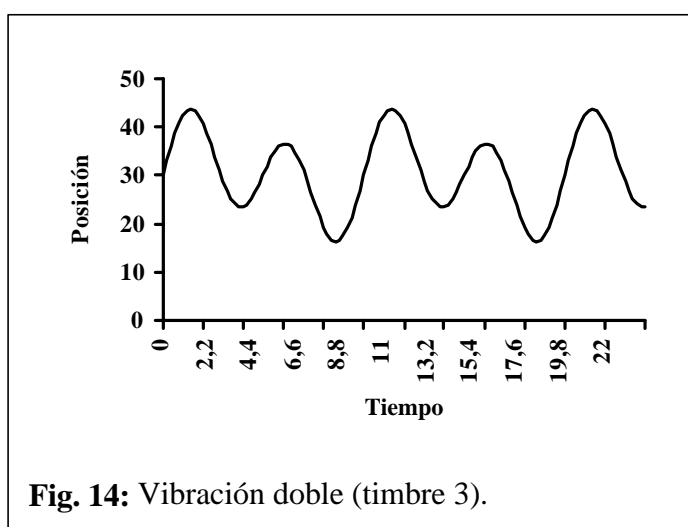
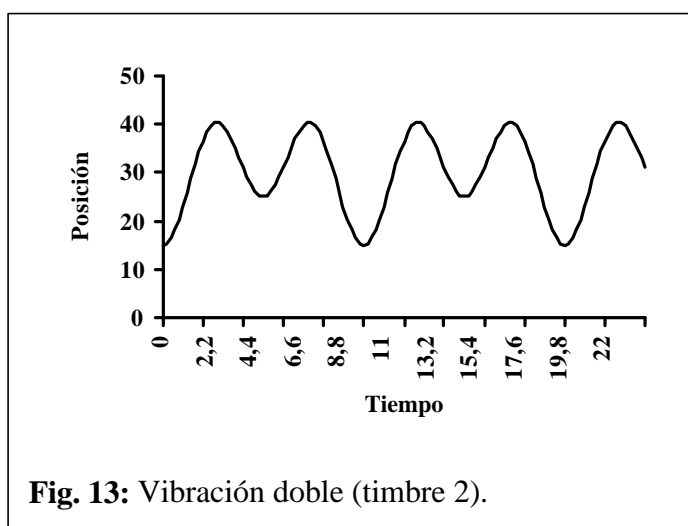
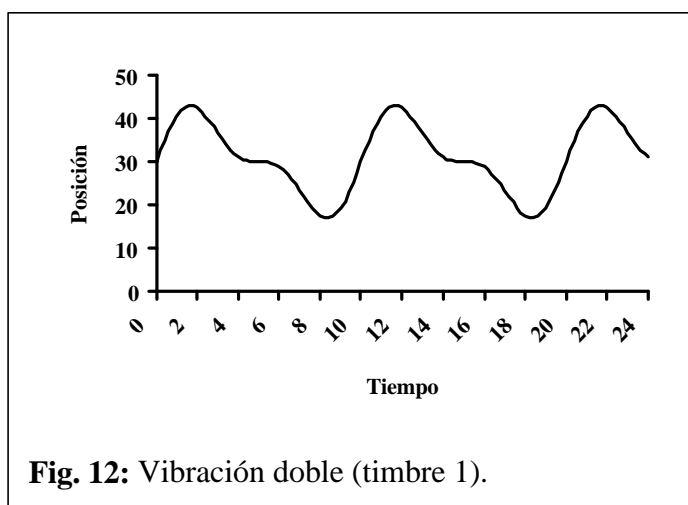
Si ahora sujetamos rígidamente la bola al marco interior (impidiendo la elongación de los muelles que la sujetan) y, por el contrario, desplazamos y soltamos el marco interior, éste (y la bola unida solidariamente a él) oscilará también con un movimiento vibratorio simple que será, en general, de diferente amplitud y frecuencia que el del caso anterior (figura 10).



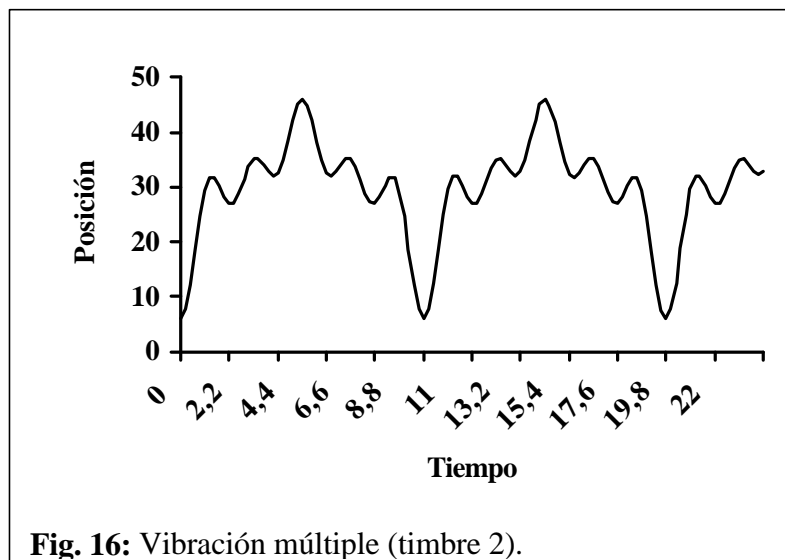
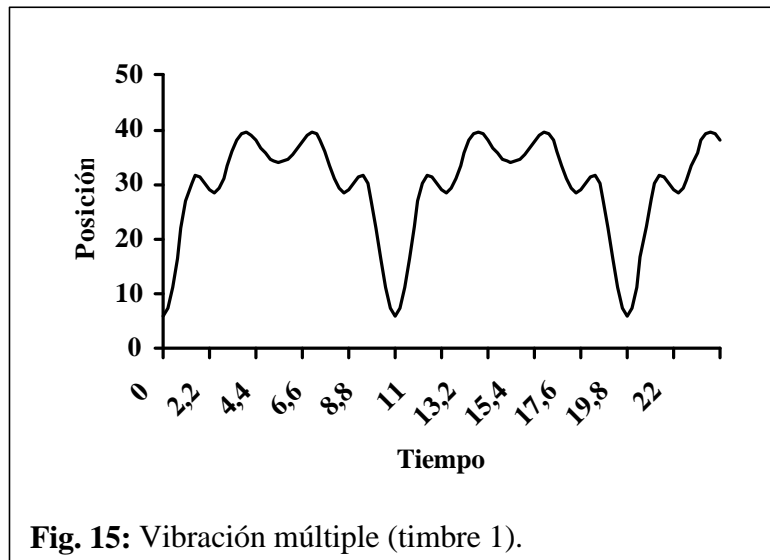
Si ahora dejamos libertad de movimiento al marco interior y a la bola, y desplazamos y soltamos el conjunto formado por ambos elementos, el movimiento de la bola con respecto al suelo podemos caracterizarlo, en primera aproximación, como la suma de los dos movimientos anteriores (figura 11). En esta gráfica podemos observar cómo la vibración sigue teniendo una amplitud (intensidad) y una frecuencia (tono) ya que el movimiento sigue siendo periódico. Pero, sin embargo, la forma de la representación no presenta la *simplicidad* del caso anterior.

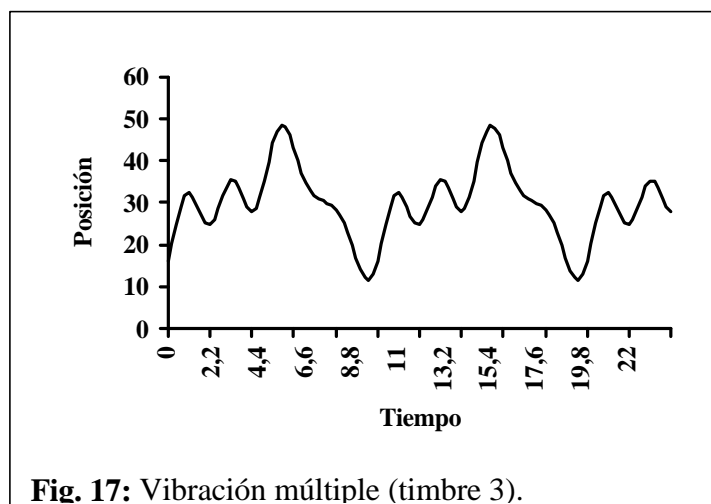


Las figuras 12, 13 y 14, representan tres vibraciones de la misma intensidad y tono pero con formas diferentes. A cada una de estas diferentes formas es a lo que se denomina *timbre*. Cada uno de dichos gráficos representa un timbre diferente.



El ejemplo de la bola sujeta con muelles puede seguir complicándose tanto como se quiera, haciendo que el segundo marco esté a su vez sujeto por muelles a un tercer marco, y éste a un cuarto, etc. Se pueden de esta forma obtener vibraciones muy complejas como las representadas en las figuras 15,16 y 17.





2.5 El sonido como vibración temporal compleja

La presentación realizada en los apartados anteriores nos faculta para poder caracterizar un sonido. Ya vimos que la fuente de un sonido era la vibración de un elemento material, como la bola de nuestro ejemplo anterior. En la mayoría de los casos, los sonidos están producidos por una vibración material compleja por lo que admite una representación temporal como las de la figura 17. En particular, los sonidos del habla, son producidos como veremos por vibraciones complejas que se representan temporalmente de la forma antedicha. Las figuras 18 y 19 presentan la representación temporal de diversos sonidos de habla.



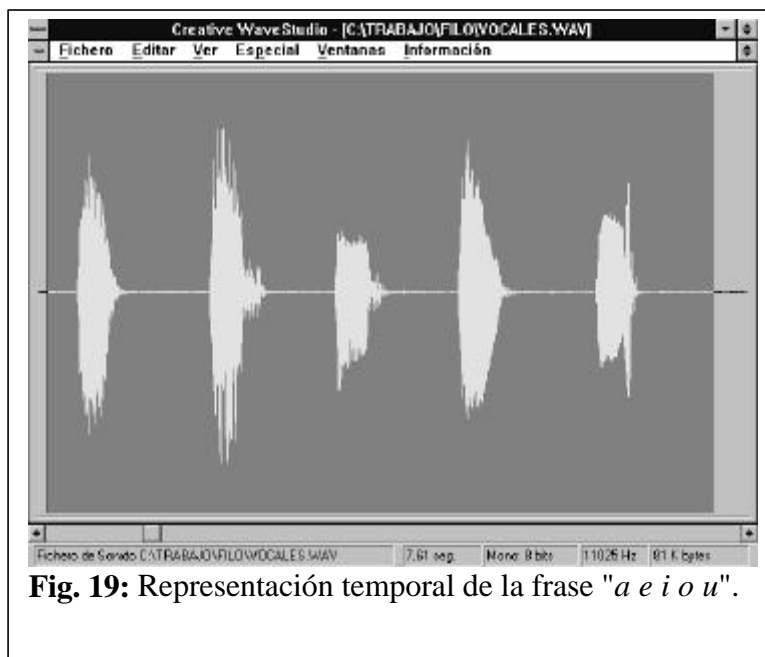


Fig. 19: Representación temporal de la frase "a e i o u".

3 CARACTERIZACIÓN ESPECTRAL DEL SONIDO

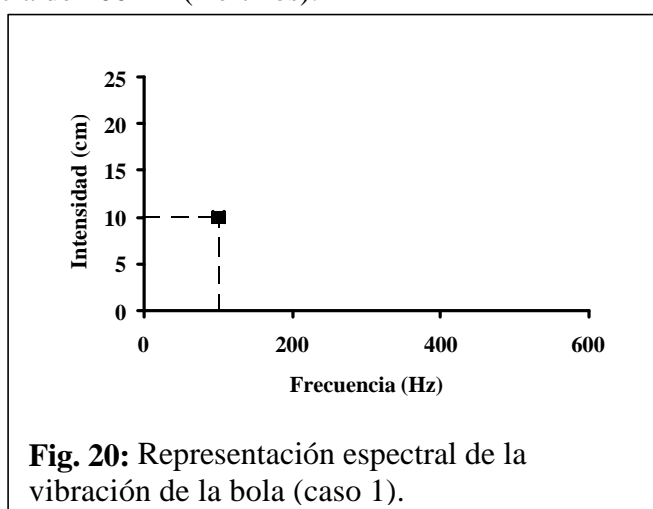
3.1 Espectro frente a tiempo

La representación temporal de los sonidos podría parecernos que es la forma *natural* de conceptualizar la realidad sonora. Una percepción en el tiempo se identifica con una representación de una onda temporal con diferentes parámetros en los que hemos reconocido las aprehensiones de intensidad, tono y timbre. A mi modo de ver, como ya apunté más arriba, la pretendida *naturalidad* de esta conceptualización deriva, principalmente, de la proximidad al carácter temporal con el que aprehendemos la realidad sonora, a lo cuál habría que sumar también que este tipo de representación es la habitual en nuestra formación académica.

Sin embargo caben otras formas de caracterizar un sonido que son tan válidas como la representación temporal. Describen la misma realidad de otra forma, pero nuestro grado de comprensión de la misma puede ser tan alto o más que en el caso anterior. La falta de familiaridad con las formas alternativas de representación no son un índice de su incapacidad sino de su novedad. Entre estas formas alternativas, hay una que es ampliamente utilizada: la representación espectral de los sonidos. Un sonido se conceptualiza como una vibración que no se describe temporalmente sino espectralmente. Pero esta representación es equivalente a la descripción en el tiempo, se refiere a la misma realidad y nos desvela en el mismo grado lo que la realidad sonora sea *en realidad*. Sin embargo esta caracterización permite una mejor manipulación de los sonidos para ciertos propósitos y, por ello, se la prefiere a la representación temporal. En concreto la representación espectral del habla nos va a permitir identificar más fácilmente su contenido: nos acerca al reconocimiento del habla.

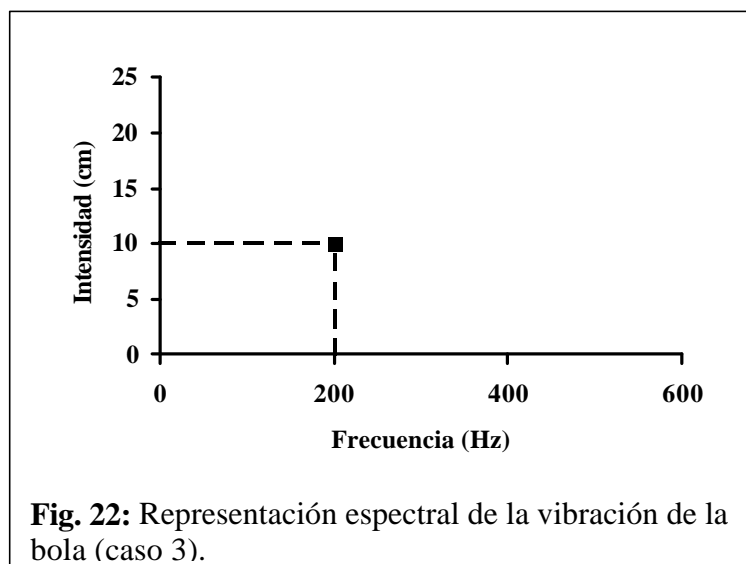
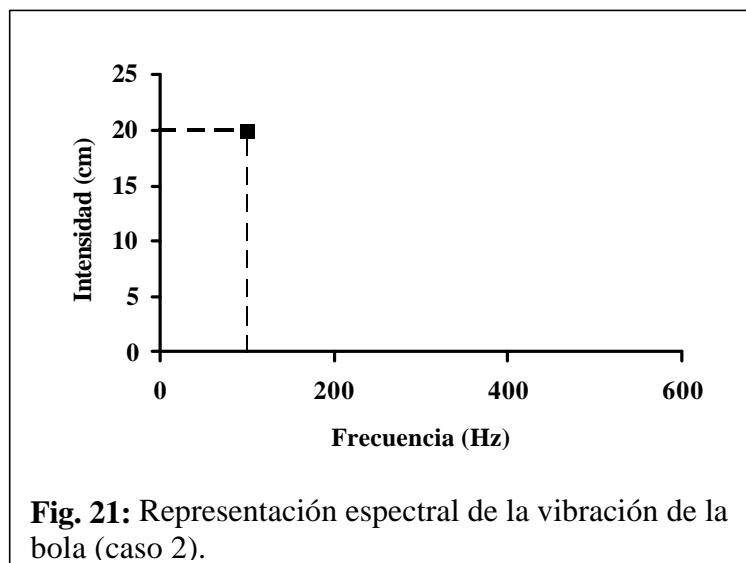
3.2 Análisis espectral de una vibración simple

Pero vayamos por pasos. Volvamos al ejemplo de la figura 1 en el que teníamos una bola sujeta con muelles a un marco de madera, y éste último sujeto al suelo. Hagamos vibrar de nuevo la bola, desplazándola 10 centímetros de su posición de equilibrio y soltándola. Pero ahora no nos interesa, en primera instancia, la posición que tome la bola. Por el contrario fijémonos simplemente en la **máxima** desviación con respecto a la posición de equilibrio. Observamos que esa desviación máxima es precisamente de 10 centímetros. A continuación nos fijamos en la frecuencia de la vibración. Para ello, a simple vista o filmándola a cámara lenta, contamos el número de vibraciones por segundo, por ejemplo contando el número de veces que la bola pasa, en un segundo, por el punto más alto. Haciendo esto nos damos cuenta que la bola va y viene 100 veces en un segundo, es decir, vibra a una frecuencia de 100 Hz (Hertzios).



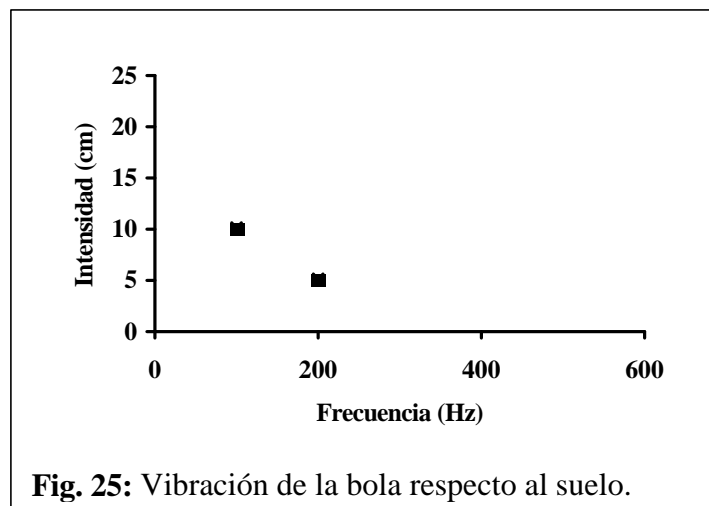
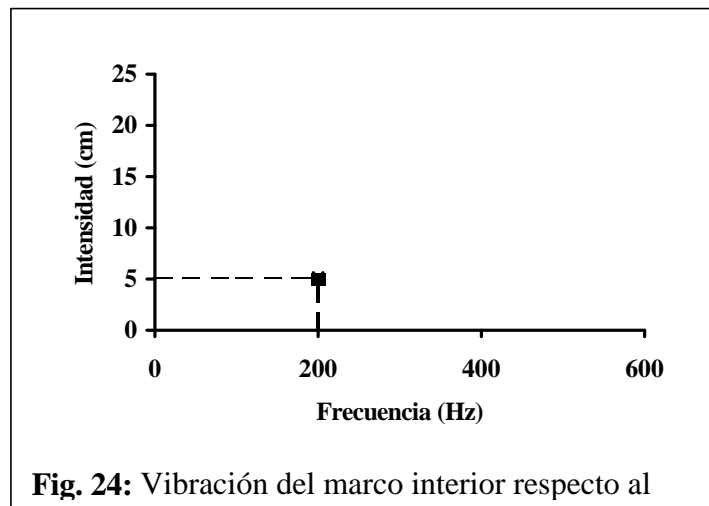
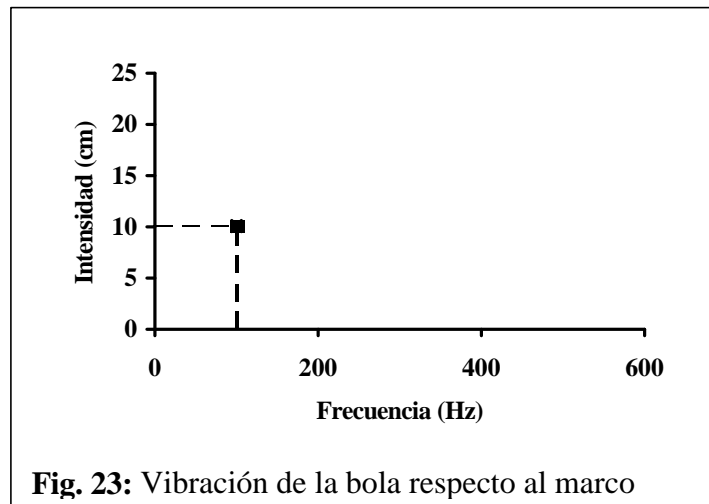
Los resultados anteriores los podemos representar gráficamente de la forma que aparece en la figura 20. Esta gráfica se denomina el espectro de la vibración (o del sonido). En ella se recoge la intensidad de la vibración y su frecuencia. Obsérvese que esta representación da la misma información que la representación temporal de la figura 5, y de hecho se puede realizar el ejercicio, muy simple por otra parte, de trazar la representación temporal dada la espectral o viceversa. Describen la misma realidad con el mismo grado de aproximación. Por tanto son equivalentes. Sin embargo, a simple vista puede comprobarse cómo la representación espectral es más simple. De un vistazo nos dice la intensidad y frecuencia de la vibración que son los parámetros pertinentes en un sonido. Es pues, para el estudio de los sonidos y del habla, una herramienta más adecuada, más fácil de interpretar. Y si esto es así en el caso de una vibración simple, ello quedará aún más claro en el caso de vibraciones o sonidos complejos.

A modo de ilustración, las figuras 21 y 22 representan espectralmente vibraciones de diferente intensidad o frecuencia que se producen al desplazar la bola 20 centímetros en lugar de 10 (figura 21), o instalar muelles más rígidos que provocan una vibración de 200 Hz en vez de la original de 100 Hz (figura 22).



3.3 Análisis espectral de una vibración compleja

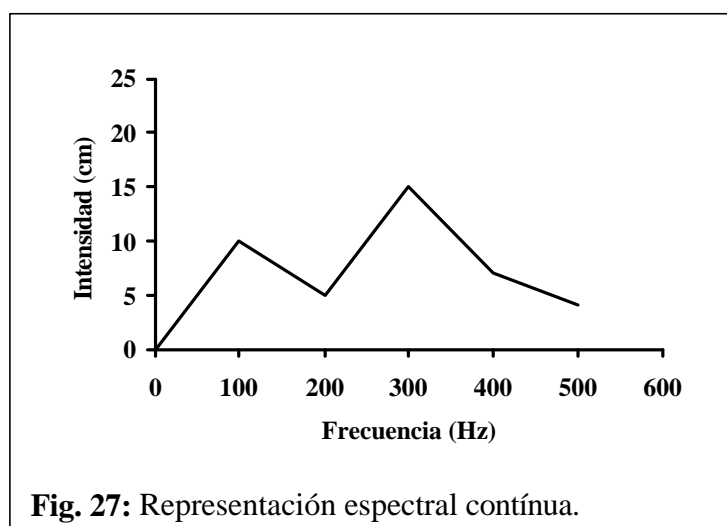
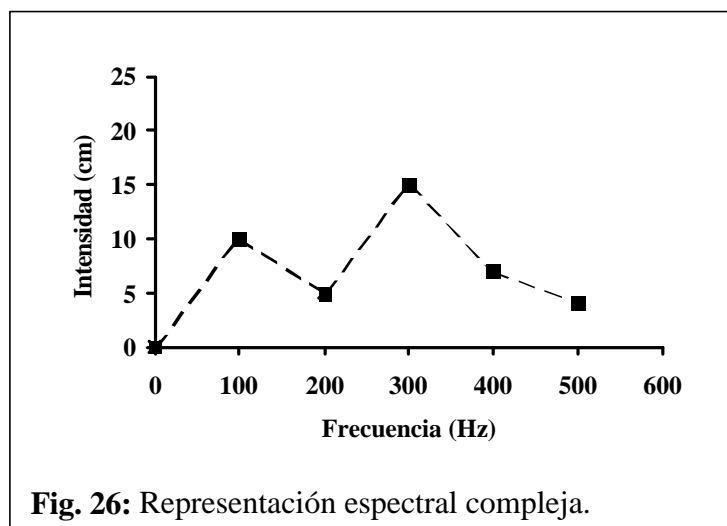
Hemos visto la caracterización espectral de una vibración simple. Estamos pues en condiciones de avanzar un paso más y considerar el caso de vibraciones complejas. Fijémonos para ello de nuevo en el artilugio de la bola sujeta a dos marcos que se representa en la figura 8. Sujetemos el marco interior y hagamos vibrar la bola. Ésta lo hará, como en el apartado anterior, con un desplazamiento máximo de 10 centímetros y una frecuencia de 100 veces por segundo. Unamos ahora solidariamente la bola al marco interior y hagamos vibrar dicho marco. En este caso supongamos que el conjunto formado por la bola y el marco se desplazan 5 centímetros como máximo de su posición de equilibrio y realizan 200 movimientos de vaivén por segundo. Las figuras 23 y 24 recogen la representación espectral de ambas vibraciones simples.



Si ahora se liberan la bola y el marco interior y se deja vibrar al artilugio en toda su complejidad, el movimiento resultante, que es la suma de los dos anteriores, se representa espectralmente mediante el gráfico de la figura 25.

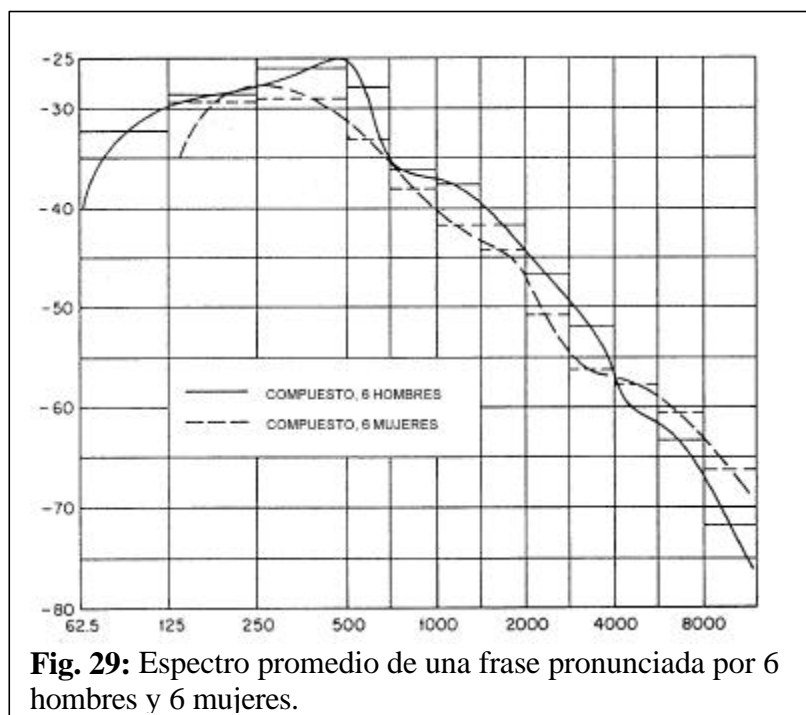
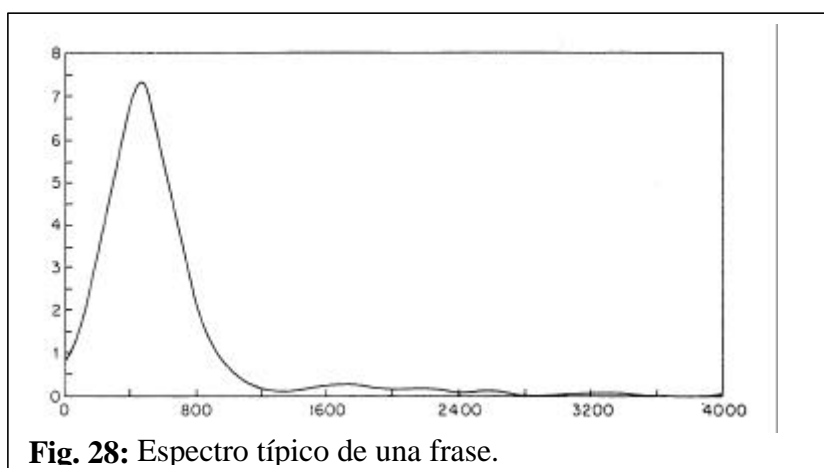
Como puede verse, esta representación es mucho más simple y fácil de entender que la equivalente representación temporal de la figura 11. Ésta y no otra es la razón de preferir el análisis espectral frente al temporal en múltiples aplicaciones de vibraciones, y en concreto en el estudio de los sonidos. A cada una de las vibraciones simples se la denomina componente espectral o *armónico*.

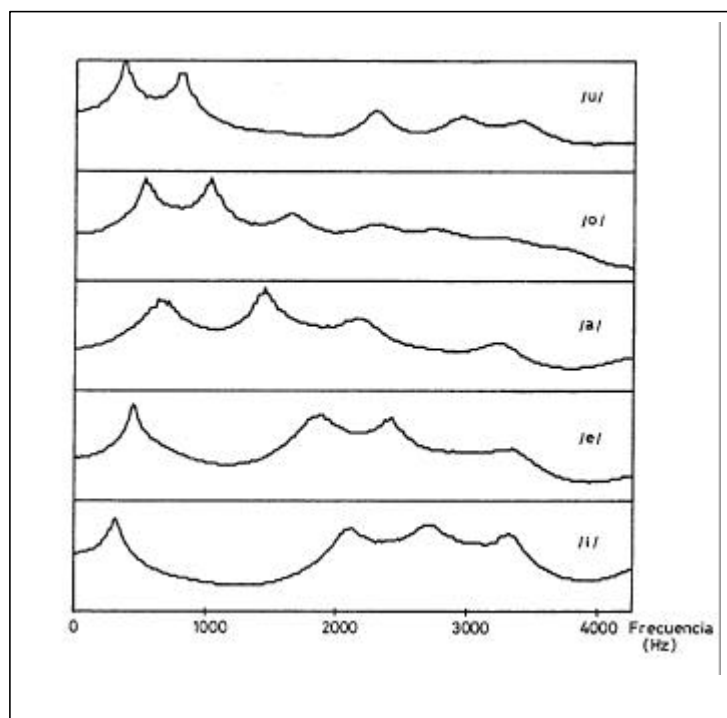
Si el artilugio se complica con una secuencia de marcos inscritos (mediante muelles) unos en otros, la vibración global resultante toma la forma, por ejemplo, de la figura 26 o de la figura 27.



3.4 El sonido como vibración espectral compleja

Un sonido en general, y uno de habla en particular, puede considerarse como el resultado de una vibración compleja en el que las *gotas* de aire hacen el papel de la bola de nuestro ejemplo anterior. Ese aire está sometido a complejos movimientos de vaivén provocados por el aparato fonador humano. La gota de aire es impulsada por los pulmones, choca, roza y rebota en la traquea, pasa por las cuerdas vocales en vibración (lo que le añade nuevas vibraciones), pasa por la faringe, se ve forzada a pasar entre la lengua, las cavidades bucal y nasal, los dientes, los labios y, por fin, alcanza el exterior. Este largo y sinuoso trayecto confiere a nuestra gota de aire un movimiento complejo que es suma de múltiples vibraciones simples. Por tanto, la representación espectral de un sonido será el de la suma de sus vibraciones simples, como ocurre en la figura 27. Las figuras 28, 29 y 30 representan espectralmente diversos sonidos de habla.





En estos gráficos puede observarse la presencia de grupos o conjuntos de armónicos de una cierta importancia relativa frente a otras bandas de frecuencia donde apenas existen componentes espectrales. Así por ejemplo en la representación espectral del fonema correspondiente a la vocal *a* (figura 30) puede observarse un grupo de armónicos en torno a los 700 Hz., otro sobre los 1.500 Hz., y otro sobre los 2.200 Hz. Cada uno de estos grupos recibe el nombre de *formante* y tienen una gran importancia para analizar el lenguaje oral. Igualmente se puede comprobar, por ejemplo, cómo el espectro de la *e* (figura 30) tiene sus tres primeros formantes en las frecuencias de 500, 1.900 y 2.400 Hz.

3.5 Espectros variables

La representación espectral pura, tal como la hemos descrito más arriba es útil para la representación y manipulación de fonemas simples. Pero a lo largo de un discurso oral el aparato fonador va articulando distintas configuraciones, va cambiando de tamaño y forma. Las gotas de aire que al principio vibraban de una manera determinada, de pronto, a causa del cambio articulatorio, comienzan a vibrar de otra forma. El espectro del sonido va pues cambiando a medida que se van articulando los diferentes fonemas. La forma de representar este hecho es mediante el uso de espectros variables en el tiempo.

Para obtener un espectro variable hay que fijarse en la forma de vibrar la gota de aire en un momento determinado, descubriendo los armónicos de la vibración. Un instante después se vuelve a analizar la vibración y se calculan sus armónicos, que pueden ser los mismo o distintos que en el caso anterior. Este proceso se va repitiendo una y otra vez, calculando el espectro de la vibración en diferentes instantes de tiempo. Todo ello se puede representar en un gráfico en tres dimensiones como los de la figura 31 y 32. En ellos se recogen la intensidad de cada armónico de frecuencia a medida que avanza el tiempo.

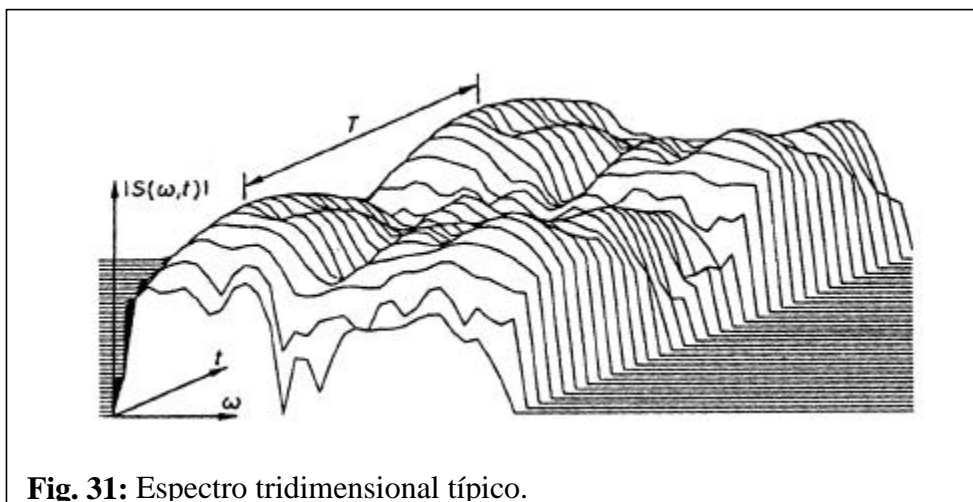


Fig. 31: Espectro tridimensional típico.

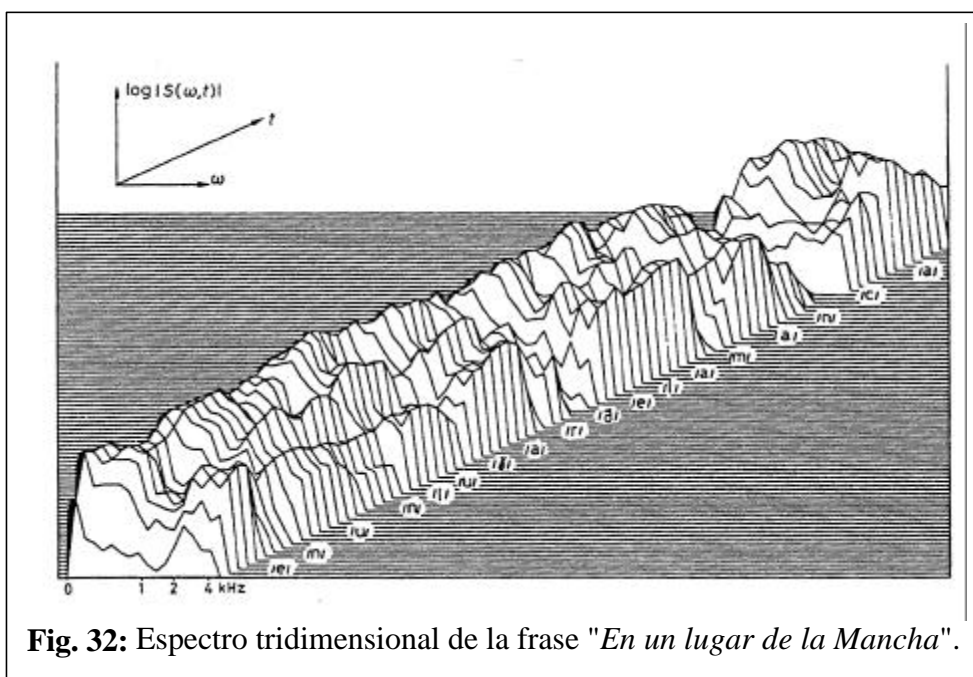


Fig. 32: Espectro tridimensional de la frase "En un lugar de la Mancha".

No obstante ese tipo de representación en tres dimensiones tiene el inconveniente de ser difícil de representar en el plano del papel. Para simplificarlo se suele recurrir en el análisis del habla a otro tipo de representación denominada *sonograma* o *espectrograma*. En un sonograma se representan en vertical las frecuencias de vibración y en horizontal el tiempo. Para un instante de tiempo dado, se representa en vertical un punto en cada componente armónico. La intensidad de ese punto es variable (más o menos negro), siendo más oscuro mientras mayor sea el valor del armónico correspondiente. Un gráfico de este tipo puede verse en la figura 33.

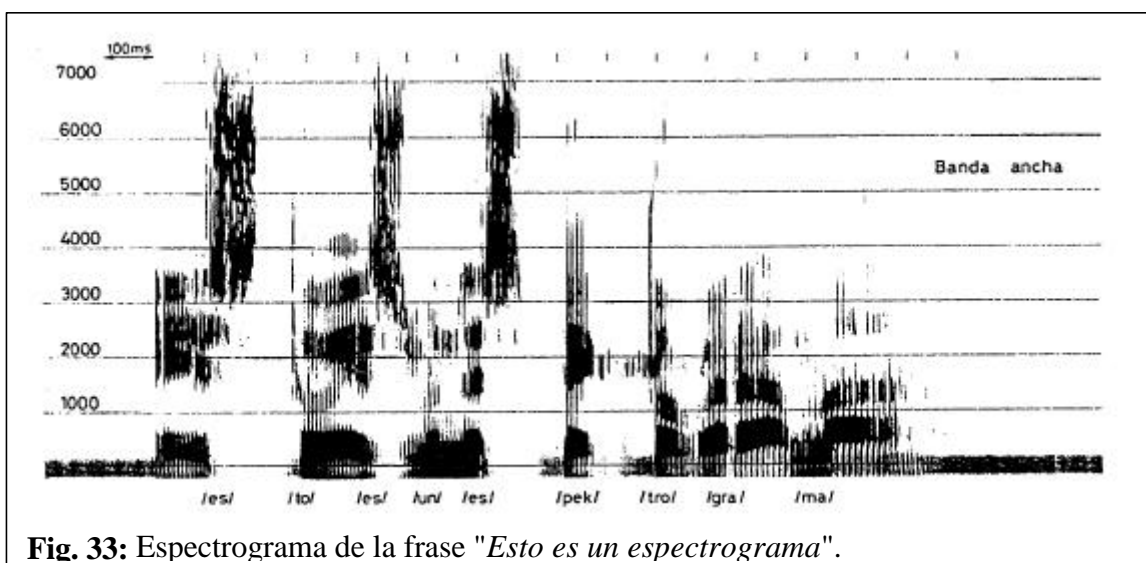


Fig. 33: Espectrograma de la frase "Esto es un espectrograma".

El sonograma nos permite, un análisis de como varía el espectro sonoro a medida que pasa el tiempo. Si tenemos en cuenta que el espectro en un momento determinado viene dado por la configuración articulatoria del aparato fonador humano, entenderemos fácilmente cómo el sonograma nos permite identificar cada una de las configuraciones articulatorias y, por tanto, cada uno de los fonemas emitidos.

4 BIBLIOGRAFÍA

- [CASA 87] Francisco Casacuberta y Enrique Vidal: *Reconocimiento Automático del Habla*. Editorial Marcombo. Barcelona, 1987.
- [FALL 85] Frank Fallside y William A. Woods, Editores: *Computer Speech Processing*. Editorial Prentice-Hall. Londres, 1985.
- [NAVA 67] T. Navarro Tomás: *Pronunciación Española*. Publicaciones de la Revista de Filología Española. Madrid, 1967.
- [QUIL 81] Antonio Quilis: *Fonética Acústica de la Lengua Española*. Editorial Gredos. Madrid, 1981.
- [RABI 78] L.R. Rabiner y R.W. Schafer: *Digital Processing of Speech Signals*. Editorial Prentice-Hall. New Jersey, 1978.