

## Representaciones discretas en tiempo–frecuencia y el problema de la selección de frecuencias \*

Alin Andrei Cârsteanu Manițiu

### Resumen

El artículo presenta la problemática general de la obtención de bases en  $\mathbf{R}^n$  con significado de frecuencia. También se presenta el efecto sobre la característica de contraste de una base, de la minimización de una función objetivo aditiva (tal como lo sería una función de costo), definida sobre los elementos de aquella base. Se discuten varios algoritmos de optimización, basados en el paquete de ondículas. Finalmente, se propone una solución novedosa a lo que es el problema del cálculo de un espectro energético significativo de la representación de una señal en una base obtenida a partir del paquete de ondículas.

*2000 Mathematics Subject Classification: 42C40, 65T60.*

*Keywords and phrases: ondícula, base ortonormal, representación tiempo–frecuencia, análisis armónico, minimización discreta, contraste.*

## 1 Introducción

El análisis en tiempo–frecuencia se dedica a la representación de señales, con soporte contiguo o discreto, en términos de un diccionario de formas de onda u ondículas. Otros términos usados en español son “onduleta” y “ondeleta”, imitaciones onomatopéyicas del francés “ondelette”. Preferimos “ondícula”, con el cual se guarda el sentido original de la palabra. Las “señales” pueden ser desde funciones en el sentido clásico hasta distribuciones y procesos estocásticos. También llamado análisis

---

\*Artículo invitado.

armónico (por razones históricas, dado que las primeras funciones analizadoras fueron funciones armónicas, siendo éste el caso de las transformadas de Fourier), el análisis en tiempo–frecuencia tiene sus raíces en la relación complementaria que guardan el tiempo y la frecuencia. Dicha relación también llama la atención a la localización en tiempo–frecuencia de los elementos de la representación, aspecto que se discutirá a continuación, así como a la posibilidad de obtener algoritmos para el uso de funciones analizadoras de formas arbitrarias (en vista que en la transformada de Fourier, el algoritmo hace uso de las propiedades de la exponencial compleja, de las cuales resulta la simetría de la transformada inversa con respecto a la directa). Así, dado el significado de amplitudes correspondientes a diferentes frecuencias, que adquieren los coeficientes de las representaciones en tiempo–frecuencia, varias formas de las funciones analizadoras tienen un interés particular, de acuerdo a su interpretación física, habiéndose compilado hasta diccionarios de funciones analizadoras. También, el análisis en tiempo–frecuencia se extiende de manera natural a soportes multidimensionales, que pueden ser asociados, en aplicaciones, con espacios físicos bidimensionales o tridimensionales, espacio–tiempos u otras entidades. El presente artículo se limitará, de acuerdo a su título, al soporte unidimensional discreto.

Consideremos una función  $f$  definida sobre un soporte discreto, formado por  $n$  puntos equidistantes:  $\mathbf{f} = [f(\tau), f(2\tau), f(3\tau), \dots, f(n\tau)]$ , donde  $n$  es una potencia de 2. En las aplicaciones, tal  $\mathbf{f}$  resulta generalmente de un muestreo. El propósito declarado del análisis armónico es de encontrar una descomposición de  $\mathbf{f}$  de la forma

$$(1) \quad \mathbf{f} = \sum_{k=1}^n c_k \mathbf{w}_k,$$

donde las  $\mathbf{w}_k = [w_k(\tau), w_k(2\tau), \dots, w_k(n\tau)]$  son  $n$  ondículas (funciones analizadoras) de una familia y los  $c_k$  son los coeficientes (o amplitudes) correspondientes a cada ondícula. Si las ondículas se escogen de tal manera que los coeficientes  $c_k$  sean únicamente determinados, entonces la familia de ondículas constituye una base en  $\mathbf{R}^n$  ( $\mathbf{f}$  es un vector real  $n$ -dimensional). Las ondículas se llaman “átomos” en tiempo frecuencia, en el sentido de que unas son versiones re–escaladas y/o trasladadas de las otras, ocupando por convención distintos rectángulos, de áreas iguales, en el plano tiempo–frecuencia:

$$(2) \quad w_k(t) = \frac{1}{\sqrt{a_k}} W\left(\frac{t - b_k}{a_k}\right),$$

donde  $W$  es la ondícula–base, los  $a_k$  son factores de escala y los  $b_k$  son factores de posición en su argumento.

En el presente artículo se usarán ondículas de Haar-Walsh [3, 9], por sus cualidades de localización, presentadas en la siguiente sección. A continuación se presentará el algoritmo de paquetes de ondículas, que permite construir varias bases de ondículas a partir de una ondícula de origen. Finalmente, se mostrarán los algoritmos que permiten escoger una base óptima con respecto a una función objetivo definida sobre el conjunto de los coeficientes, así como el contenido en frecuencias de dichas bases y algunas de sus aplicaciones.

## 2 Paquetes de ondículas

Fieles al propósito de descomponer la señal  $\mathbf{f}$  en función de las frecuencias presentes en ella, consideremos dos operadores, un operador  $A$  de “alta frecuencia” y un operador  $B$  de “baja frecuencia”, definidos sobre  $\mathbf{R}^2$  de la siguiente manera:

$$(3) \quad \begin{cases} A[x, y] = (x - y)/\sqrt{2} \\ B[x, y] = (x + y)/\sqrt{2}. \end{cases}$$

Observemos que al aplicar los operadores  $A$  y  $B$  a una señal  $\mathbf{f}$  resulta  $[A, B]\mathbf{f} = [A\mathbf{f}, B\mathbf{f}] = [A[f_1, f_2], \dots, A[f_{n-1}, f_n], B[f_1, f_2], \dots, B[f_{n-1}, f_n]]$  y se obtiene una base en  $\mathbf{R}^n$  (recordemos que  $n$  es divisible por 2, siendo una potencia de 2), dado que los operadores son lineales e invertibles, de manera que

$$(4) \quad \begin{cases} (A[x, y] + B[x, y])/\sqrt{2} = x \\ (A[x, y] - B[x, y])/\sqrt{2} = y. \end{cases}$$

Es fácil verificar que esta base es ortonormal. Observemos también que en el cambio de base, a cada par de coordenadas en la base temporal le corresponde un par de coordenadas en la nueva base. Por lo tanto, las coordenadas de cada par en la nueva base cubren exactamente el doble de la extensión temporal de una coordenada en la base temporal (el período del muestreo), y difieren entre sí únicamente en su contenido de frecuencia. Una representación intuitiva de la posición de los elementos de las dos bases en el plano tiempo–frecuencia, para  $n = 8$ , se muestra en la figura 1.

El sistema de funciones ortogonales (3) fue propuesto por primera vez por Haar [3] y Walsh [9].



de veces en cada elemento. Por lo tanto, llamaremos a las bases con esta propiedad “bases fundamentales”. En el caso más general en que  $n$  es una potencia cualquiera de 2, hay  $\log_2(n) + 1$  bases fundamentales, correspondientes al aplicar los operadores  $0, 1, \dots, \log_2(n)$  veces. Notemos que cada base fundamental es exhaustiva con respecto a las  $2^k$  permutaciones posibles (con repeticiones) de los dos operadores  $A$  y/o  $B$ , aplicados  $k$  veces, sobre un subconjunto de  $n/2^k$  puntos del dominio de definición de  $\mathbf{f}$ . Por lo tanto, el conjunto de bases fundamentales contiene todos los elementos de bases que se puedan obtener aplicando los operadores  $A$  y/o  $B$ ; llamamos a este conjunto el “paquete de ondículas”.

Una observación importante para la obtención de diferentes bases usando los operadores  $A$  y  $B$  es el hecho siguiente: debido a que la transformación (3) involucra cada vez solamente 2 elementos de una base, sobre cada subconjunto  $\{m2^k + 1, \dots, (m + 1)2^k\}$ ,  $k \in \{1, \dots, \log_2(n)\}$  y  $m \in \{0, \dots, n/2^k - 1\}$ , del dominio de definición de  $\mathbf{f}$  se pueden aplicar los operadores  $A$  o  $B$ , en cualquier orden, cualquier número de veces entre 0 y  $k$ , y el resultado sería de nuevo una base en  $\mathbf{R}^n$  (al ser lineal e invertible la transformación (3), la ortonormalidad de tales bases es menos trivial y será discutida aparte). Un ejemplo de base construida de este modo es la base–ondícula (representada en la figura 3 para  $n = 8$ ). La propiedad distintiva de esta base es que, siendo sus operadores (de altas a bajas frecuencias)  $A, AB, ABB, ABBB$  etc., en cada banda de frecuencia, la escala (el factor  $a_k$  de la ecuación (2), o equivalentemente el número de elementos de la base temporal involucrados en la representación del elemento en cuestión de la presente base) es inversamente proporcional con la frecuencia de aquella banda. Históricamente, ésta fue la primera base en tiempo–frecuencia desarrollada a parte de las bases fundamentales.

$Af_{12}$	$Af_{34}$	$Af_{56}$	$Af_{78}$
$ABf_{1-4}$		$ABf_{5-8}$	
$ABBf_{1-8}$			
$BBBf_{1-8}$			

Figura 3: Localización en el plano tiempo–frecuencia de las coordenadas en la base–ondícula.

La base frecuencial resulta, si se toma la base temporal como base estándar, después de  $\mathcal{O}(n \log(n))$  operaciones. Este orden de velocidad es similar a la del algoritmo de la transformada rápida de Fourier [2], dado que siguiendo la definición se trata de  $n$  convoluciones de la función  $f$ , definida en  $n$  puntos, con diferentes ondículas de éste mismo tamaño, tales como son representadas en la figura 4.

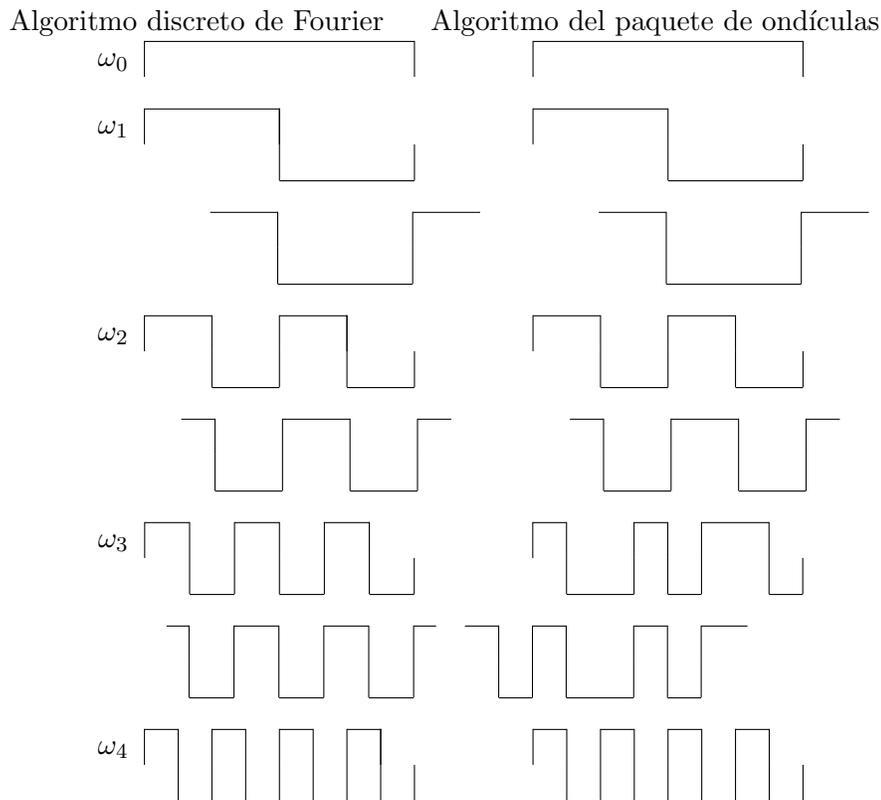


Figura 4: Formas de las ondas analizadoras correspondientes a diferentes frecuencias  $\omega$  para el algoritmo del paquete de ondículas (derecha) comparadas con las del algoritmo discreto de Fourier (izquierda), ambos algoritmos usan la función analizador de Haar.

Wickerhauser [10] demuestra la equivalencia, en primer orden (y por lo tanto, también asintóticamente, para  $n$  grande), entre el algoritmo de los paquetes de ondículas y el algoritmo de Fourier. Además, se puede

construir un paquete similar al de ondículas aplicando el algoritmo de Fourier sucesivamente a los  $n$ ,  $n/2 + n/2$ ,  $n/4 + n/4 + n/4 + n/4$ , etc. puntos del dominio de definición de  $\mathbf{f}$ . Sin embargo, el algoritmo de Fourier es “rápido” solamente para bases trigonométricas, mientras que el algoritmo presentado en esta sección mantiene el mismo  $\mathcal{O}(n \log(n))$  para cualquier par de operadores  $[A, B]$ .

Demostremos a continuación la ortonormalidad de las bases construidas a partir del paquete de ondículas Haar. (a) Por su modo de construcción, que consiste en aplicar el conjunto de operadores  $[A, B]$  a subconjuntos  $\{m2^k + 1, \dots, (m+1)2^k\}$ , para  $k \in \{1, \dots, \log_2(n)\}$  y  $m \in \{0, \dots, n/2^k - 1\}$  del dominio de definición de  $\mathbf{f}$ , dos elementos arbitrarios de una base obtenida a partir del paquete de ondículas Haar tienen que ser ajenos o en sus representaciones por elementos de la base temporal, o en sus representaciones por elementos de la base frecuencial (o en ambos casos). Supongamos, sin pérdida de generalidad, que sea éste el caso de la base temporal, y consideremos la misma base temporal como base estándar. (Podemos hacer la suposición sin pérdida de generalidad debido a que el producto de  $A$  y  $B$  aplicados a un par de vectores unitarios  $[\mathbf{1}_k, \mathbf{1}_j]$  resulta ser  $A[\mathbf{1}_k, \mathbf{1}_j] \cdot B[\mathbf{1}_k, \mathbf{1}_j] = \frac{1}{2}\mathbf{1}_k^2 - \frac{1}{2}\mathbf{1}_j^2 = 0$ .) Entonces los dos elementos son trivialmente ortogonales. (b) La normalidad de las bases está garantizada por el hecho que los operadores  $[A, B]$  conservan las normas en  $L_2$ :  $(A[f_k, f_j])^2 + (B[f_k, f_j])^2 = \frac{1}{2}(f_k + f_j)^2 + \frac{1}{2}(f_k - f_j)^2 = f_k^2 + f_j^2$ .

Notemos que la ortonormalidad de los operadores  $A, B$  no es suficiente para la ortonormalidad de toda base así obtenida del paquete de ondículas; sino que es necesario que los elementos involucrados en cada aplicación de los operadores sean los mismos como en el caso de los operadores Haar-Walsh. De lo contrario, al aplicar la transformación de base, por intermedio de tales operadores, a un número de  $n$  elementos, su norma llega a encontrarse, después de la transformación, en un número de elementos mayor que  $n$ , o bien al restringirnos a  $n$  elementos, la norma de éstos disminuiría generalmente con respecto a la norma inicial. Estos efectos son llamados efectos terminales, y si no son muy importantes en magnitud, se dice que las respectivas bases son casi ortonormales. Varios autores (por ejemplo, Herley et al. [5]) proponen para estos casos la ortonormalización de las bases, por procedimientos tipo Gramm-Schmidt o similares. Notemos que con este proceder se pierde el significado en tiempo–frecuencia de dichas bases. (Mientras que este significado no es importante, por ejemplo, para fines de com-

presión de señales, para la mayoría de las aplicaciones de investigación sí lo es.)

### 3 Optimalidad de las bases construidas del paquete de ondículas

#### 3.1 Criterios globales y funciones objetivo locales

El propósito de esta optimización es de escoger la base que minimiza una función objetivo, definida sobre los coeficientes de la representación de  $\mathbf{f}$  en la base, de tal manera que se optimicen ciertas características cualitativas globales de dicha representación. Tal característica es, por ejemplo, el llamado “contraste”. Intuitivamente, aumentar el contraste quiere decir aumentar los elementos mayores y disminuir los elementos menores. En términos de una base que resultaría en la representación más contrastante de un vector dado, se elige de un conjunto de bases aquella que maximiza las coordenadas mayores (en valor absoluto) del vector considerado (la ortonormalidad de las bases asegura la minimización de las coordenadas menores). En las aplicaciones, maximizar el contraste es útil para fines de compresión de señales, haciendo posible la eliminación de un mayor número de elementos pequeños (compresión aproximada) o ceros (compresión exacta), así como para fines de investigación en la estructura en tiempo–frecuencia de las señales, revelando las frecuencias dominantes por un período limitado de tiempo.

Una manera para definir el contraste en una representación es la de ordenar las coordenadas de manera decreciente y construir la secuencia de normas (usuales, en  $L_2$ ) de los vectores formados por las primeras  $1, 2, \dots, n$  coordenadas, lo que es equivalente (para bases ortonormales) con la secuencia de sumas parciales de los cuadrados de las primeras  $1, 2, \dots, n$  coordenadas. Por su construcción, esta secuencia representa una función no convexa, llamémosla  $\sigma(k)$  en función del número  $k$  de coordenadas involucradas, haciendo la convención  $\sigma(0) = 0$ . Podemos definir la superioridad del contraste de una primera representación en comparación con una segunda representación: si  $\sigma_1(k) \geq \sigma_2(k)$ ,  $\forall k \in \{1, \dots, n\}$ , y  $\exists k \in \{1, \dots, n\}$  tal que  $\sigma_1(k) > \sigma_2(k)$ . Observemos que, en el caso general, esta definición no introduce un orden completo en el conjunto de representaciones: las gráficas de las funciones  $\sigma(k)$  de dos determinadas representaciones se pueden cruzar, en cuyo caso ninguna de las dos (de acuerdo a esta definición) tiene un contraste superior a la

otra. Un modo sencillo de resolver este problema sería de fijar un cierto  $k_0$ , para que sea el único punto en donde comparamos  $\sigma_1(k_0)$  y  $\sigma_2(k_0)$ . Para algunas aplicaciones en particular, fijar un tal  $k_0$  podría ser útil, pero aquí preferimos analizar el caso general, aun con la desventaja de no tener un orden completo de las representaciones por medio del contraste.

Queda entonces encontrar una función objetivo aditiva, definida sobre los coeficientes de la representación de  $\mathbf{f}$  en una base, de tal manera que se optimice el contraste. Demostraremos en este contexto el siguiente teorema:

**Teorema 3.1.1** *Sean las representaciones de una señal  $\mathbf{f}$  en dos bases ortonormales distintas, con una de las representaciones más contrastada que la otra, en el sentido que  $\sigma_1(k) \geq \sigma_2(k)$ ,  $\forall k \in \{1, \dots, n\}$ , y  $\exists k \in \{1, \dots, n\}$  tal que  $\sigma_1(k) > \sigma_2(k)$ , y sea  $s(\cdot)$  una función objetivo real, cóncava. Entonces*

$$\sum_{k=1}^n s([c_k^2]_1) < \sum_{k=1}^n s([c_k^2]_2).$$

*Demostración:* Probamos que se puede construir una operación que transforma la secuencia ordenada  $\{[c_k]_1\}$  a  $\{[c_k]_2\}$  en un número finito de pasos, tal que en cada paso ocurre una disminución de  $\sum_{k=1}^n s(c_k^2)$ . Sea

$$\delta := \min_{k \in \{l, \dots, m-1\}} [\sigma_1(k) - \sigma_2(k)]$$

tal que  $\delta > 0$ ,  $\sigma_1(l-1) = \sigma_2(l-1)$  y  $\sigma_1(m) = \sigma_2(m)$ ,  $n \geq m > l > 0$ . La existencia de una tal subsecuencia  $l, \dots, m-1$  está garantizada por  $\sigma_1(0) = \sigma_2(0) = 0$ ,  $\sigma_1(n) = \sigma_2(n)$  por ortonormalidad de las bases, y  $\exists k \in \{1, \dots, n\}$  tal que  $\sigma_1(k) > \sigma_2(k)$ . Cambiemos entonces  $[c_l^2]_2 \leftarrow [c_l^2]_2 + \delta$  y  $[c_m^2]_2 \leftarrow [c_m^2]_2 - \delta$ , lo que significa  $\sigma_2(k) \leftarrow \sigma_2(k) + \delta$ ,  $\forall k \in \{l, \dots, m-1\}$ . Como previsto, encontramos una disminución de  $\sum_{k=1}^n s([c_k^2]_2)$ :  $s([c_l^2]_2 + \delta) + s([c_m^2]_2 - \delta) < s([c_l^2]_2) + s([c_m^2]_2)$ , como  $\delta > 0$ ,  $l \geq m \iff [c_l]_2 \geq [c_m]_2$ , y la función  $s$  es cóncava. En este momento, para por lo menos un punto  $k_0 \in \{l, \dots, m-1\}$  tenemos  $\sigma_1(k_0) = \sigma_2(k_0)$ . Podemos repetir este procedimiento para subsecuencias terminando, y respectivamente empezando, en  $k_0$ . La no-conveidad y la monotonicidad de  $\sigma$  serán respetadas dentro de las subsecuencias, como solamente hay una traslación vertical, y serán también respetadas alrededor de  $k_0$ , como  $\sigma_2(k_0 - 1) \leq \sigma_1(k_0 - 1) \leq \sigma_1(k_0) =$

$\sigma_2(k_0) \leq \sigma_2(k_0 + 1) \leq \sigma_1(k_0 + 1)$ . Por lo tanto, la existencia y el orden de  $[c_k]_2$  será preservado a cada paso, realizando una disminución de  $\sum_{k=1}^n s([c_k^2]_2)$  al respectivo paso siguiente, hasta que

$$\begin{aligned} \max_{k \in \{1, \dots, n\}} \{\sigma_1(k) - \sigma_2(k)\} = 0 &\iff \sigma_1(k) = \sigma_2(k), \forall k \in \{1, \dots, n\} \\ &\iff [c_k]_1 = [c_k]_2, \forall k \in \{1, \dots, n\}, \end{aligned}$$

es decir que las dos representaciones se vuelven idénticas y el algoritmo se termina (después de un número de pasos menor o igual a  $n$ ).  $\square$

### 3.2 Algoritmos de optimización

En la sección precedente hemos mostrado como escoger una función objetivo, definida sobre las coordenadas de  $\mathbf{f}$  en una base, y cuya minimización optimice el contraste de la representación de  $\mathbf{f}$  en dicha base. Queda entonces diseñar algoritmos para buscar dentro del paquete de ondículas la base ortonormal que minimice una función objetivo dada. La búsqueda está basada en el hecho que para realizar cambios de base usando los operadores  $A$  y  $B$  sobre 2 coordenadas temporalmente consecutivas de una base, se puede reemplazar este par de coordenadas con un par de coordenadas vecinas en frecuencia, y recíprocamente, como se mostró en la sección 2.

El primer algoritmo de minimización de la función objetivo, actuando solamente sobre un subconjunto de bases del paquete de ondículas, se encuentra en Wickerhauser [10], y se conoce como el algoritmo (sencillo) del árbol (binario).

Para ilustrar el algoritmo, es útil poner las representaciones en bases tiempo–frecuencia de las figuras 1 y 2 en forma del árbol de la figura 5, en el cual cada nivel es la representación de la señal en una base, correspondiendo (de abajo hacia arriba) de base temporal hasta la base

$BBB\mathbf{f}$	$ABB\mathbf{f}$	$AAB\mathbf{f}$	$BAB\mathbf{f}$	$BAA\mathbf{f}$	$AA\mathbf{f}$	$ABA\mathbf{f}$	$BBA\mathbf{f}$
$BB\mathbf{f}$		$AB\mathbf{f}$		$AA\mathbf{f}$		$BA\mathbf{f}$	
$B\mathbf{f}$				$A\mathbf{f}$			
$\mathbf{f}$							

Figura 5: Árbol del paquete de ondículas.

frecuencial. (Observemos que para  $n > 2^3$ , hay más niveles en el árbol, hacia arriba, para llegar a la base frecuencial y completar todo el paquete de ondículas.) Para escoger los elementos de la base optimizada en este caso, empezando con la base frecuencial, se compara la función objetivo para cada par de coordenadas adyacentes en frecuencia con el par que les corresponde por la ecuación (3) (es decir que se comparan los elementos correspondientes entre la penúltima y la última base – la frecuencial), escogiendo cada vez entre los dos pares aquel que minimice la función objetivo. A continuación, se comparan cuartetos de coordenadas de la misma frecuencia en la antepenúltima base con los cuartetos correspondientes de la base obtenida al paso anterior, escogiendo entre los cuartetos correspondientes cada vez aquel que minimice la función objetivo. Después, se comparan octetos de coordenadas de la misma frecuencia en la precedente base con los octetos correspondientes de la base obtenida al paso anterior, y así sucesivamente hasta la base temporal. El algoritmo es fácil de seguir, de arriba hacia abajo, sobre el árbol de la figura 5. Sin embargo, aunque el algoritmo minimice la función objetivo sumada sobre los elementos de la base (con las consecuencias globales discutidas en la sección precedente), se observa inmediatamente que este algoritmo no tiene acceso a todas las bases que se puedan construir a partir del paquete de ondículas; además, induce una asimetría artificial entre tiempo y frecuencia.

El mismo algoritmo puede ser empleado empezando con la base temporal (llamémoslo algoritmo inverso del árbol sencillo). Aunque ésto no mejora el algoritmo, ha sugerido la idea para otro algoritmo, operando sobre un conjunto sensiblemente más grande de bases, y publicado por Herley et al. [4, 5], así como por Ramchandran y Vetterli [7]. El llamado algoritmo doble del árbol consta en correr el algoritmo sencillo a partir de cada una de las bases fundamentales (correspondientes a cada uno de los niveles del árbol), teniendo en consecuencia  $\mathcal{O}(n \log^2(n))$  operaciones. El algoritmo guarda las mismas desventajas que el algoritmo sencillo, en el sentido que todavía no tiene acceso a todas las bases que puedan resultar del paquete de ondículas (lo que resultará obvio en la presentación de los siguientes algoritmos), y también guarda una asimetría entre tiempo y frecuencia (por correr hacia la base temporal, o posiblemente en el otro sentido, pero no en ambos).

Finalmente, algoritmos que escogen entre todas las bases posibles del paquete de ondículas fueron construidos independientemente por Thiele y Villemoes [8], Herley et al. [6] y Cârsteanu et al. [1]. Estos algoritmos están basados en la misma idea (presentada a continuación),

diferenciándose solamente en el manejo de las variables en memoria (para minimizar la memoria ocupada [1] o el tiempo de ejecución del algoritmo [8, 6], respectivamente). Observemos primero que las bases ortonormales que se pueden construir por las relaciones (3) y (4) a partir del paquete de ondículas no pueden contener simultáneamente elementos de la base temporal y de la base frecuencial (por no ser ortogonales entre sí). Esta constatación ofrece la clave para construir un algoritmo que tenga acceso a todas las bases ortonormales que se puedan construir a partir del paquete de ondículas, y de ella resulta directamente el siguiente lema:

**Lema 3.2.1** *Sea  $\mathcal{P}(\{1, \dots, n\})$  el conjunto de bases ortonormales en  $\mathbf{R}^n$  que se pueden construir a partir del respectivo paquete de ondículas,  $\mathcal{F}(\{1, \dots, n\})$  el subconjunto de  $\mathcal{P}(\{1, \dots, n\})$  de bases que no contienen ningún elemento de la base temporal, y  $\mathcal{T}(\{1, \dots, n\})$  el subconjunto de  $\mathcal{P}(\{1, \dots, n\})$  de bases que no contienen ningún elemento de la base frecuencial. Entonces*

$$\mathcal{F}(\{1, \dots, n\}) \cup \mathcal{T}(\{1, \dots, n\}) = \mathcal{P}(\{1, \dots, n\}).$$

De este lema desprendemos que cualquier base ortonormal construida a partir del paquete de ondículas (y por lo tanto la base óptima buscada) pertenece a  $\mathcal{F}(\{1, \dots, n\})$  o a  $\mathcal{T}(\{1, \dots, n\})$  (o a los dos), por lo cual la búsqueda algorítmica de la base óptima dentro de  $\mathcal{P}(\{1, \dots, n\})$  se reduce a dos búsquedas dentro de  $\mathcal{F}(\{1, \dots, n\})$  y de  $\mathcal{T}(\{1, \dots, n\})$ , respectivamente. Obviamente, para  $n > 2$ , tenemos  $\mathcal{F}(\{1, \dots, n\}) \cap \mathcal{T}(\{1, \dots, n\}) \neq \emptyset$ , lo que significa que se puede mejorar el algoritmo de búsqueda en términos de rapidez (a precio de usar mas memoria) memorizando  $\mathcal{F}(\{1, \dots, n\}) \cap \mathcal{T}(\{1, \dots, n\})$  y buscando una sola vez en este subconjunto (en lugar de dos veces, al buscar dentro de  $\mathcal{F}(\{1, \dots, n\})$  y dentro de  $\mathcal{T}(\{1, \dots, n\})$ ). Thiele y Villemoes [8] muestran que memorizando todas las comparaciones hechas en cada etapa, el algoritmo llega a tener  $\mathcal{O}(n \log(n))$  operaciones. Veremos a continuación dos lemas, que nos permitirán continuar la búsqueda de la base óptima dentro de  $\mathcal{T}(\{1, \dots, n\})$  y dentro de  $\mathcal{F}(\{1, \dots, n\})$ , respectivamente:

**Lema 3.2.2** *Las bases del subconjunto  $\mathcal{T}(\{1, \dots, n\})$  contienen solamente elementos de las primeras  $n - 1$  bases del respectivo paquete. Por lo tanto, tenemos que  $\mathcal{P}(\{1, \dots, n/2\}) \times \mathcal{P}(\{n/2 + 1, \dots, n\}) = \mathcal{T}(\{1, \dots, n\})$ . Es decir que  $\mathcal{T}(\{1, \dots, n\})$  consta del paquete construido de los primeros  $n/2$  elementos de la base temporal y del paquete construido de los últimos  $n/2$  elementos de dicha base.*

**Lema 3.2.3** *Las bases del subconjunto  $\mathcal{F}(\{1, \dots, n\})$  contienen solamente elementos de las últimas  $n-1$  bases del respectivo paquete. Por lo tanto,  $\mathcal{F}(\{1, \dots, n\})$  consta del paquete construido de los primeros  $n/2$  elementos de la base temporal y del paquete construido de los últimos  $n/2$  elementos de dicha base.*

Hemos entonces reducido, en 2 pasos intermedios, la búsqueda de la base óptima dentro de  $\mathcal{P}(\{1, \dots, n\})$  a la búsqueda dentro de 4 paquetes de bases de dimensión  $n/2$ . Como el conjunto generado por los 4 paquetes es idéntico a  $\mathcal{P}(\{1, \dots, n\})$  (cualquier base que exista en  $\mathcal{P}(\{1, \dots, n\})$  se encuentra también en el conjunto), concluimos que el algoritmo toma en cuenta todas las bases ortonormales que se puedan construir del paquete de bases de dimensión  $n$  y, por inducción, que el algoritmo encuentra el resultado después de  $\mathcal{O}(n^2)$  pasos (sin tomar en cuenta ningún modo de eliminar las redundancias entre  $\mathcal{F}(\cdot)$  y  $\mathcal{T}(\cdot)$  mencionadas anteriormente).

Observemos sin embargo que  $\#(\mathcal{P}(\{1, \dots, n\}))$ , el número de bases contenidas en  $\mathcal{P}(\{1, \dots, n\})$ , no crece con  $n$  de manera polinomial, sino exponencial, lo que haría un procedimiento por comparaciones directas prácticamente imposible de usar. Efectivamente, como por razones de simetría tenemos  $\#(\mathcal{P}(\{1, \dots, n\})) = 2\#^2(\mathcal{P}(\{1, \dots, n/2\})) - \#^4(\mathcal{P}(\{1, \dots, n/4\}))$ , y como  $\#(\mathcal{P}(\{1\})) = 1$  y  $\#(\mathcal{P}(\{1, 2\})) = 2$ , tenemos que  $\exists \beta$  tal que  $\lim_{n \rightarrow \infty} \#(\mathcal{P}(\{1, \dots, n\}))/\beta^n = (\sqrt{5} - 1)/2$ , con  $\beta \approx 1.84454757\dots$  [1]. Para tener una idea del número de bases que pueden tomar en cuenta los algoritmos antes presentados, mencionemos que para el algoritmo del árbol doble,  $\beta \approx 1.7148445\dots$ , mientras que para el algoritmo del árbol sencillo,  $\beta \approx 1.5028368\dots$  (el límite es igual a 1 en los dos últimos casos) [8]. Para precisar esta comparación visualizando gráficamente ciertas diferencias entre los algoritmos, en la figura 6 se presentan (para  $n = 8$ ) las localizaciones en tiempo–frecuencia de: (A) una base accesible a los tres algoritmos; (B) una base que no puede ser tomada en cuenta por el algoritmo del árbol sencillo; y (C) una base a la cual no tienen acceso ni el algoritmo del árbol sencillo, ni el algoritmo del árbol doble.

## 4 Aplicaciones

Como ya discutido, maximizar el contraste de una señal es útil para fines de compresión, haciendo posible la eliminación de un mayor número de

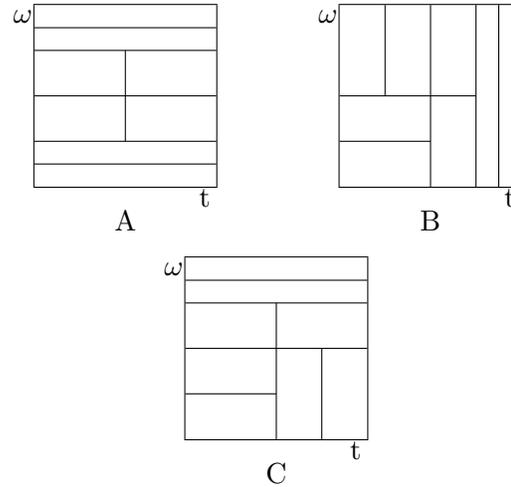


Figura 6: Representaciones en tiempo–frecuencia de: (A) una base accesible a los tres algoritmos; (B) una base que no puede ser tomada en cuenta por el algoritmo del árbol sencillo; y (C) una base a la cual no tienen acceso ni el algoritmo del árbol sencillo, ni el algoritmo del árbol doble.

elementos pequeños para compresión aproximada, o ceros para compresión exacta. En la figura 7, presentamos la base temporal, así como la base óptima de una señal de 256 muestras. La función objetivo escogida es la entropía, la señal llegando de una entropía de 5.1444... en la base temporal a una entropía de 1.1858... en la base óptima. Dado que la entropía es una función cóncava, se realiza también un contraste mas fuerte en la base óptima (figura 8). Este contraste nos permite reconstruir aproximadamente la señal usando solamente 20 coordenadas de la base óptima (figura 9).

Para fines de investigación en la estructura en tiempo–frecuencia de una señal, una representación óptima puede revelar las frecuencias dominantes por un período limitado de tiempo, lo que no es realizable ni en la representación temporal, ni en la representación frecuencial. La figura 10 nos muestra la representación en su base óptima de un ejemplo de una tal señal con frecuencia variable en el tiempo, un chirrido.

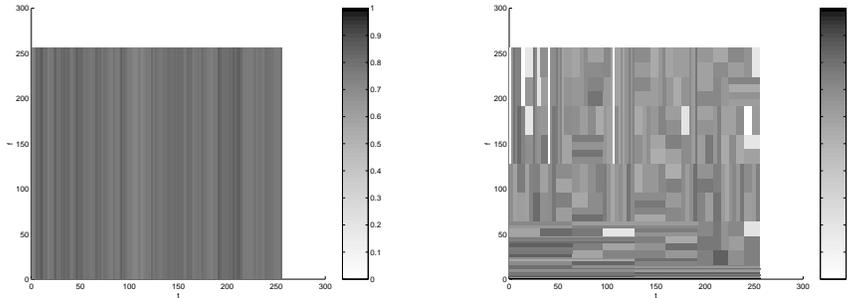


Figura 7: Representación de una señal de 256 muestras en la base temporal (izquierda), así como la base óptima (derecha): las amplitudes (coordenadas en las respectivas bases) están representadas por intensidad de gris.

## 5 El problema de las frecuencias para el espectro energético

En las secciones precedentes hemos visto cómo, a partir del paquete de ondículas, se pueden construir bases en  $\mathbf{R}^n$ , con significado en tiempo–frecuencia, que minimicen una función aditiva sobre los elementos de la base, y en consecuencia adquieren ciertas características de contraste. Sin embargo, una vista más cuidadosa al paquete de ondículas revela que, igual que en el caso del algoritmo (real) discreto de Fourier, los pares de elementos consecutivos en frecuencia (con excepción del primero y del último elemento) de cualquier base construida, son caracterizados en realidad por la misma frecuencia, y difieren solamente por una traslación ( $b_k$  de la ecuación (2)). Se presenta entonces el problema de cómo obtener la norma de las coordenadas correspondientes a cada frecuencia, el conjunto de dichas normas constituye el llamado “espectro energético”. Dicho espectro, interpretado como “contenido” de energía en cada frecuencia (también “densidad” de energía en el caso de espectros continuos) tiene varias aplicaciones en procesos físicos caracterizados por frecuencias. Para ejemplificar, el espectro energético de una señal que representa la variación en el tiempo de la velocidad en un punto de un fluido turbulento es gobernado por las leyes de escalamiento que rigen el mismo fluido, y son el resultado de la transferencia de energía dentro del rango inercial de escalas del fluido, desde las escalas de

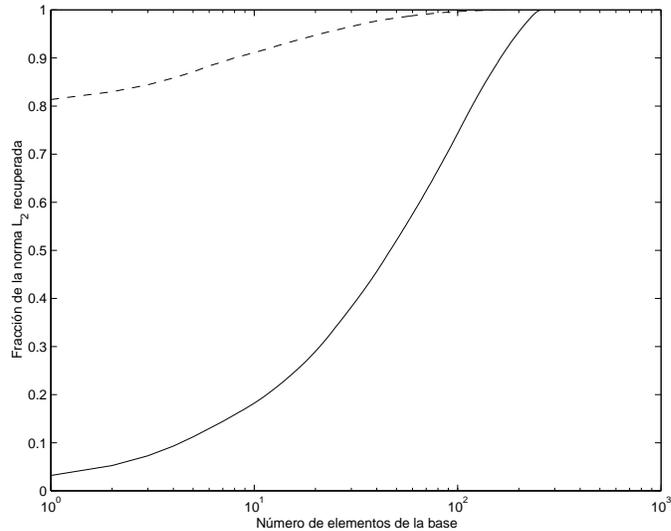


Figura 8: Fracción de la norma  $L_2$  de la señal reconstruida, en función del número de las coordenadas mayores usadas en la reconstrucción: notemos la superioridad del contraste de la base óptima (línea interrumpida) comparado con el contraste de la base temporal (línea continua).

inyección de energía hacia el rango viscoso. El estudio del respectivo espectro energético ha permitido explicar, aunque todavía no completamente, el fenómeno de movimiento turbulento. Por lo tanto, estimamos como muy importante la posibilidad de extraer el espectro energético de una señal a partir de las bases construidas en tiempo–frecuencia.

Para poder evaluar todas las normas correspondientes a frecuencias en una base, en ella cada elemento de la base temporal debe entrar en el cálculo de elementos de una sola representación fundamental. El conjunto de bases que satisfacen esta condición es precisamente el conjunto de bases accesibles por el algoritmo inverso del árbol sencillo (recordemos que por “inverso” hemos entendido el intercambio de tiempo con frecuencia en el algoritmo usual). Esto, porque en el algoritmo del árbol sencillo, la descomposición frecuencial es constante en el tiempo [10], y por lo tanto, en el algoritmo inverso, la descomposición temporal es la misma para todas las frecuencias. Lo que quiere también decir que para

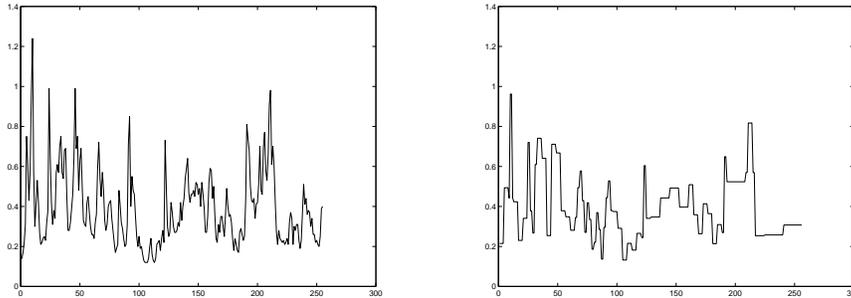


Figura 9: La señal original de 256 muestras (izquierda) y la señal reconstruida de 20 coordenadas de la base óptima (derecha).

las bases obtenidas por los demás algoritmos, en general no se puede calcular un espectro energético con significado.

Proponemos aquí una visión diferente acerca de este asunto: calcular el espectro energético con solamente  $n/2 + 2$  en lugar de  $n$  elementos. El siguiente lema nos ayuda a hacerlo:

**Lema 5.1.1** *Usando el algoritmo para obtener la base óptima, si paramos la descomposición en frecuencia a 4 elementos en lugar de 2, y si anulamos para propósitos de comparación los 2 elementos extremos de los 4 (aparte del caso que son los mismos extremos de una base fundamental), procuramos comparar todas las bases construidas de esta manera, a partir del paquete de ondículas.*

Para probarlo, es suficiente observar que el algoritmo es equivalente al aplicar el algoritmo usual al conjunto  $\mathcal{F}$  descrito en la sección anterior.

La ventaja de este acercamiento radica en el hecho que para la mayoría de las señales, las  $n/2 + 2$  coordenadas mayores de la base óptima contienen casi toda la energía de la señal, mientras que la base óptima tiene en la mayoría de los casos un  $\sigma(k)$  superior a lo de la base obtenida por el algoritmo inverso del árbol sencillo, para todo  $k \in \{1, \dots, n/2 + 2\}$ .

## 6 Conclusiones

Hemos presentado aquí una revisión del estado actual de los algoritmos de optimización discreta para representaciones en tiempo–frecuencia,

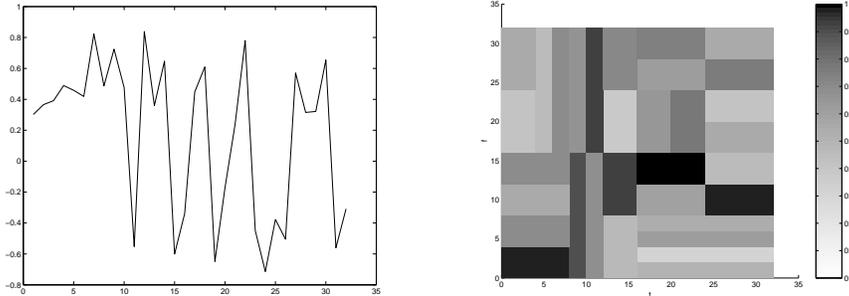


Figura 10: Un chirrido y su representación en la base óptima (donde las coordenadas están representadas por intensidad de gris): observemos cómo las frecuencias dominantes son delimitadas durante diversos intervalos de tiempo.

enfocando el trabajo en ejemplificar las características especiales y los logros de cada algoritmo. Una atención particular fue extendida al problema de la selección de frecuencias, la cual juega un papel importante en la definición del espectro energético de una señal. En este contexto, observamos que una versión nueva de un algoritmo existente puede resolver el problema, y también proponemos un algoritmo nuevo, basado en un acercamiento diferente, que surge de los lemas comprobados en el presente artículo. Si bien este algoritmo resulta superior en términos de contraste, la cuestión de la optimización del contraste entre todas las bases que permiten la definición de un espectro energético queda como problema abierto para el futuro.

### Agradecimientos

El autor quiere agradecer a los editores de Morfismos la invitación a escribir el presente artículo, así como a los tres revisores anónimos quienes, con sus comentarios, contribuyeron a mejorar el texto, y a Jesús González Espino Barros por su revisión final.

Alin Cârsteanu Manițiu  
*Departamento de Matemáticas,*  
 CINVESTAV-IPN  
 A.P. 14-740  
 México D.F. 07000  
 México  
 alin@math.cinvestav.mx

## Referencias

- [1] Cârsteanu A.; Sapozhnikov V. B.; Venugopal V.; Fofoula-Georgiou E., *Absolute optimal time-frequency basis – a research tool*, J. Phys. A: Math. Gen., **30** (1997), 7133-7146.
- [2] Cooley J. W.; Tukey J. W., *An algorithm for machine calculation of complex Fourier series*, Math. Computation, **19** (1965), 297-301.
- [3] Haar A., *Zur Theorie der orthogonalen Funktionensysteme*, Math. Ann., **69** (1910), 331-371.
- [4] Herley C.; Kovačević J.; Ramchandran K.; Vetterli M., *Arbitrary orthogonal tilings of the time-frequency plane* (Int. Symp. on Time-Frequency and Time-Scale Analysis, 11-14), Victoria, BC 1992.
- [5] Herley C.; Kovačević J.; Ramchandran K.; Vetterli M., *Tilings of the time-frequency plane: construction of arbitrary orthogonal bases and fast tiling algorithms*, IEEE Trans. Signal Processing, **41** (1993), 3341-3359.
- [6] Herley C.; Xiong Z.; Ramchandran K.; Orchard M. T., *Joint space-frequency segmentation using balanced wavelet packet trees for least cost image representation*, IEEE Trans. Image Processing, **6** (1997), 1213-1230.
- [7] Ramchandran K.; Vetterli M., *Best wavelet packet bases in a rate-distortion sense*, IEEE Trans. Image Processing, **2** (1993), 160-173.
- [8] Thiele C. M.; Villemoes L. F., *A Fast Algorithm for Adapted Time-Frequency Tilings*, Applied and Computational Harmonic Analysis, **3** (1996), 91-99.
- [9] Walsh J. L., *A Closed Set of Normal Orthogonal Functions*, Amer. J. Math., **45** (1923), 5-24.
- [10] Wickerhauser M. V., *Lectures on wavelet packet algorithms*, Preprint, Department of Mathematics, Washington University, 1991.