

Titulo: Estudios estadísticos de partituras musicales

Las partituras musicales son una representación escrita precisa de una pieza musical. Gracias a esto, se tiene conocimiento de la música que se ha compuesto de los últimos 500 años.

La estructura de la música es un sistema complejo sometido a factores culturales, acústicos, técnicas de interpretación y percepción de la audiencia. La perspectiva de la física estadística en la música ha sido muy interesante en las últimas décadas. Desde la identificación de patrones temporales o leyes de potencia, hasta el desarrollo de algoritmos de composición musical.

En años recientes se han llevado a cabo en el Instituto de Ciencias Físicas de la UNAM y el Centro de Investigación en ciencias de la UAEM, estudios estadísticos de partituras musicales [1, 2]. Se ha abordado el tema considerando aspectos de no linealidad, irreversibilidad, asimetría y relaciones de escalamiento.

En 2017, González-Espinoza, Larralde, Martínez-Mekler, & Müller et al., publicaron un análisis de piezas musicales de compositores de épocas que abarcan desde el barroco del siglo XVI, hasta la música contemporánea del siglo XX, En este análisis se investigó relaciones de escalamiento y presencia de características No Lineales en partituras musicales.

Para encontrar relaciones de escalamiento se utilizó el método DFA, que permite la detección de la autosimilitud intrínseca incrustada en una serie temporal que puede ser no estacionaria y evita la detección espuria de una autosimilitud aparente, que puede ser un artefacto de tendencias extrínsecas. [3,4]

Primero, se obtiene la función acumulada de la serie temporal (de longitud total N), de la forma:

$$y(j) = \sum_{i=1}^j [x(i) - \langle x \rangle] \quad (1)$$

Donde $\langle x \rangle$ es el valor promedio de los datos $x(i)$. De esta forma, en el caso de que la serie de tiempo original esté acotada, esto no afectará en la exhibición de relaciones de escalamiento y pueda ser cuantificado por un parámetro de autosimilitud. [4]

La serie de tiempo acumulada es dividida en ventanas de la misma longitud s , donde en cada una se realizará un ajuste polinomial de grado m , denotado por $y_{m,s}(j)$. La siguiente etapa es eliminar la tendencia de la serie temporal acumulada, restándolo con el ajuste local en cada ventana. Entonces, dado una ventana de tamaño s , la magnitud característica de la fluctuación para esta serie temporal acumulada y sin tendencia se calcula mediante la media cuadrática:

$$F(s) = \sqrt{\frac{1}{N} \sum_{j=1}^N [y(j) - y_{m,s}(j)]^2} \quad (2)$$

Este cálculo se repite para diferentes escalas de tiempo (tamaño de ventana s) para proveer una relación entre $F(s)$ y el tamaño de la ventana s . Si las fluctuaciones en pequeñas ventanas están relacionadas con las fluctuaciones en ventanas más largas en una forma de ley de potencia, entonces, las fluctuaciones se pueden caracterizar mediante un exponente de escala. La pendiente del ajuste lineal entre $\log[F(s)]$ y $\log[s]$ determina el exponente de escalamiento (parámetro de autosimilitud) α .

El parámetro α de autosimilitud tiene una relación con la función de autocorrelación $C(\tau)$ de la señal original de la forma $C(\tau) \sim \tau^{2\alpha-2}$, donde α es el exponente de Hurst H cuando está en un rango de $0 < H \leq 1$. Un H mayor que 0.5 y menor o igual que 1, indica correlaciones persistentes de ley de potencias de largo alcance. Cuando $0 < H < 0.5$, anticorrelaciones de ley de potencias están presentes. Cuando $H = 0.5$, entonces corresponde a ruido blanco donde un valor a cierto instante está completamente decorrelacionado con previos y futuros valores; y cuando $\alpha > 1$, entonces la relación entre el parámetro de autosimilitud y el exponente de Hurst es de la forma $H = \alpha - 1$.

También se toma en cuenta que el espectro de potencias $S(f)$ de la señal original también es de la forma de ley de potencia $S(f) \sim 1/f^\beta$. Y como consecuencia del teorema de Wiener-Khinchin, que establece que la transformada de Fourier de la función de autocorrelación es el espectro de potencias de la señal, α está relacionado de la forma $\beta = 2\alpha - 1$. El caso de $\alpha = 1$ corresponde al ruido rosa $1/f$.

Estos estudios destacan la complejidad inherente en la estructura musical, por lo que nos pueden llevar a reflexionar sobre el determinismo y su relación con la no linealidad. Si consideramos una relación lineal, estamos hablando de una relación de proporcionalidad, representando un determinismo claro. Sin embargo, cuando hablamos de sistemas no lineales, su relación con el determinismo no es tan claro, puesto que existen sistemas no lineales deterministas, como las ecuaciones de Lorenz; y sistemas no lineales no deterministas, como las fluctuaciones en el mercado de valores.

El objetivo de esta tesis es complementar los estudios estadísticos previos sobre partituras musicales anteriores al siglo XX, enfocándose en investigar aspectos deterministas en estos datos empíricos. La teoría del caos y los sistemas dinámicos no lineales podrían aplicarse para analizar las secuencias melódicas, descubriendo si ciertas melodías siguen reglas deterministas complejas que no son inmediatamente evidentes.

Tradicionalmente, detectar propiedades deterministas no líneas en series de tiempo empíricos se basa en la reconstrucción de espacio fase, ya que en este espacio es donde tales propiedades se manifiestan. Sin embargo, esto solo es aplicable para datos estacionarios, en gran medida libres de ruido, de un sistema de baja dimensión y requiere el ajuste no trivial de varios parámetros.

El método clásico para identificar propiedades deterministas de un sistema es haciendo una reconstrucción del espacio fase a través del procedimiento del embedding. Para explicar el embedding, consideremos el siguiente ejemplo: supóngase que queremos estudiar un sistema dado. Para ello, capturamos una serie de datos en función del tiempo, lo que se conoce como una serie de tiempo. No obstante, nuestro sistema tiene un espacio fase real que nosotros desconocemos, el cual es necesario conocer para poder estudiar nuestro sistema.

Para abordar esta limitación, utilizamos nuestra serie de tiempo y lo ploteamos contra sí misma, pero con un retraso en el tiempo. Podemos utilizar tantas dimensiones como sea necesario con retrasos apropiados. De esta manera, reconstruimos el espacio fase de nuestro sistema, un proceso conocido como dinámica embebida. Sin embargo, para obtener una

representación adecuada del espacio fase, es crucial seleccionar el retraso correcto y utilizar un número suficientes de dimensiones.

Es importante destacar que las dinámicas embebidas son difeomorfos, es decir, no hay cambios en la topología en comparación con el espacio fase original, esta característica pertenece al teorema de Takens.

El retraso τ correcto debe cumplir con la condición que sea mayor que cero para garantizar que las trayectorias en el espacio reconstruido no se crucen, y también debe ser lo suficientemente grande para permitir que las diferencias en las trayectorias sean perceptibles para la computadora.

Asimismo, el número de dimensiones utilizado debe ser mayor que el doble del número de dimensiones del espacio fase original para establecer un embedding real, es decir, uno que sea topológicamente correcto. Sin embargo, determinar estas condiciones para un sistema real que se desea estudiar puede resultar desafiante, además, es una condición poco estricta. Por ejemplo, se ha demostrado que un buen embedding del sistema de Lorenz se logra con tan solo 3 dimensiones, a pesar de que el sistema en sí mismo es de 3 dimensiones. En teoría, se requerirían 7 dimensiones para que el embedding funcionara correctamente.

La motivación principal de este proyecto de tesis es probar un nuevo método para detectar características deterministas en datos empíricos, desarrollado por investigadores en colaboración entre el ICF UNAM y CinC UAEM. Este método presenta un índice basado en las fases de Fourier que detecta el grado de irregularidad o comportamiento azaroso de series temporales con un nivel de significancia bien definido, extrayendo sus propiedades no lineales y evitando la engorrosa reconstrucción del espacio fase.

El índice J se basa en la transformada de Fourier de una señal real $S(t)$

$$\mathcal{F}[S(t)] = A(f)e^{i\phi(f)}$$

Donde $A(f)$ son las amplitudes de Fourier y sus correspondientes fases de Fourier se denotan como $\phi(f)$. Dado que la transformada de Fourier es biyectiva, la información codificada en la serie temporal está igualmente presente en su transformada de Fourier.

Las amplitudes de Fourier contienen únicamente las propiedades univariantes lineales, mientras que las fases de Fourier contienen las propiedades no lineales. Las fases de Fourier son las más interesantes porque es la no linealidad que proporciona la más grande cantidad de información, desde las características deterministas y no lineales, hasta interrelaciones con otras señales.

Aunque una de las dificultades de extraer información a través de las fases de Fourier es que no presentan una estructura clara y fácil de interpretar, en su lugar, son difusas y aparentemente una distribución de puntos aleatorios.

Es aquí donde el índice J nos ayudará a interpretar las fases de Fourier y retornar un número que nos ayudará a reconocer características deterministas en un nivel de significancia bien definido.

A partir de nuestra serie de tiempo que representa el tono de la melodía principal de la pieza musical seleccionada, se aplica la transformada de Fourier para obtener las fases de Fourier correspondientes. Como estamos trabajando un caso univariante, definimos vectores bidimensionales de las fases, considerando las señales retardadas $S(t)$ y $S(t + \tau)$.

Como las fases $\phi(f)$ están definidas sobre un círculo unitario, están restringidas entre $[0, 2\pi]$, lo que significa que la secuencia de los vectores bidimensionales $[\phi_1(f_i), \phi_2(f_i)]$ definen la trayectoria en la superficie de un toro de 2 dimensiones.

En la superficie del toro, al moverse de un punto $[\phi_1(f_i), \phi_2(f_i)]$ a otro $[\phi_1(f_{i+1}), \phi_2(f_{i+1})]$, el caminante tiene la flexibilidad de desplazarse tanto hacia adelante como hacia atrás en las direcciones toroidal y poloidal. La distancia entre estos dos puntos puede ser más grande o más chica dependiendo del camino que se elija sobre la superficie del toro. Por lo que consideramos 9 posibilidades de caminos en el toro para determinar el más corto. Estos caminos pueden ser $[\phi_1(f) + 2\pi, \phi_2(f)], [\phi_1(f), \phi_2(f) + 2\pi], [\phi_1(f) - 2\pi, \phi_2(f)], [\phi_1(f), \phi_2(f) - 2\pi], [\phi_1(f) + 2\pi, \phi_2(f) + 2\pi], \dots$

Elegir el camino más corto cumple con la definición de métrica, eso es, que debe ser definida positiva, simétrica y debe obedecer a la desigualdad del triángulo.

A continuación, se mide el ángulo de desviación α en el espacio que cubre el toro utilizando la métrica euclidiana. Estimamos posibles correlaciones dentro de la trayectoria sobre el toro utilizando el siguiente índice estadístico J en el círculo unitario:

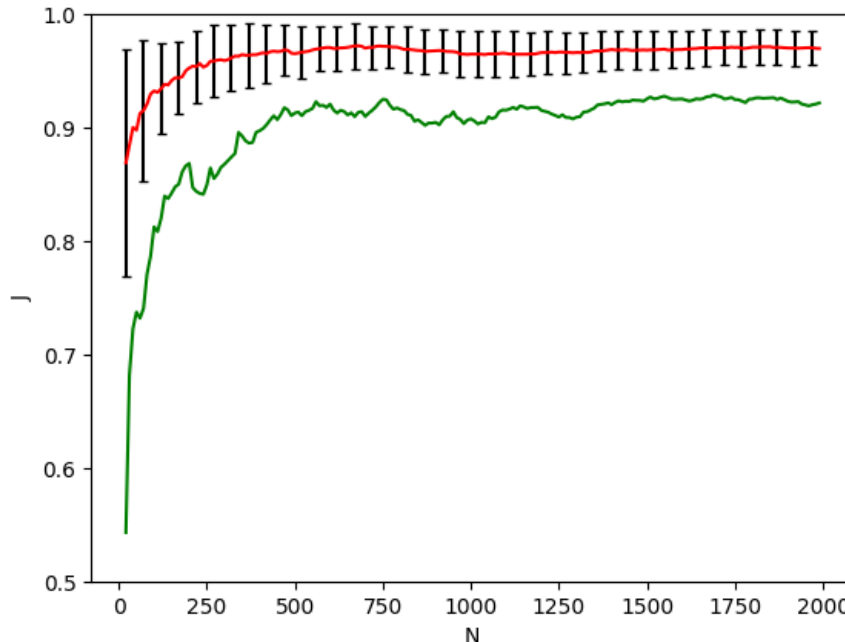
$$J = 1 - \left| \frac{1}{N-1} \sum_{l=1}^{N-1} e^{i\alpha(f_l)} \right|$$

Este índice toma valores entre cero, que representa un ángulo de desviación constante, y uno, que representa una trayectoria completamente aleatoria.

Para poder comparar estadísticamente cualquier resultado, es necesario tener datos que representen la hipótesis nula. Para esto, generamos secuencias de fases aleatorias bidimensionales $[\varphi_1(f), \varphi_2(f)]$ del mismo tamaño N . Donde $\varphi(f)$ son números independientes aleatorios uniformemente distribuidas entre 0 y 2π .

Para este trabajo, se generaron 100 secuencias de fases aleatorias. Para conservar interrelaciones lineales potenciales entre las dos señales, la diferencia $\varphi_1 - \varphi_2$ se mantiene igual para todas las 100 secuencias.

En la gráfica se presenta la curva promediada del índice J para las 100 secuencias, con su respectiva desviación estándar. También se construyó una curva de los valores mínimo de J como función de la longitud N de la serie de tiempo.

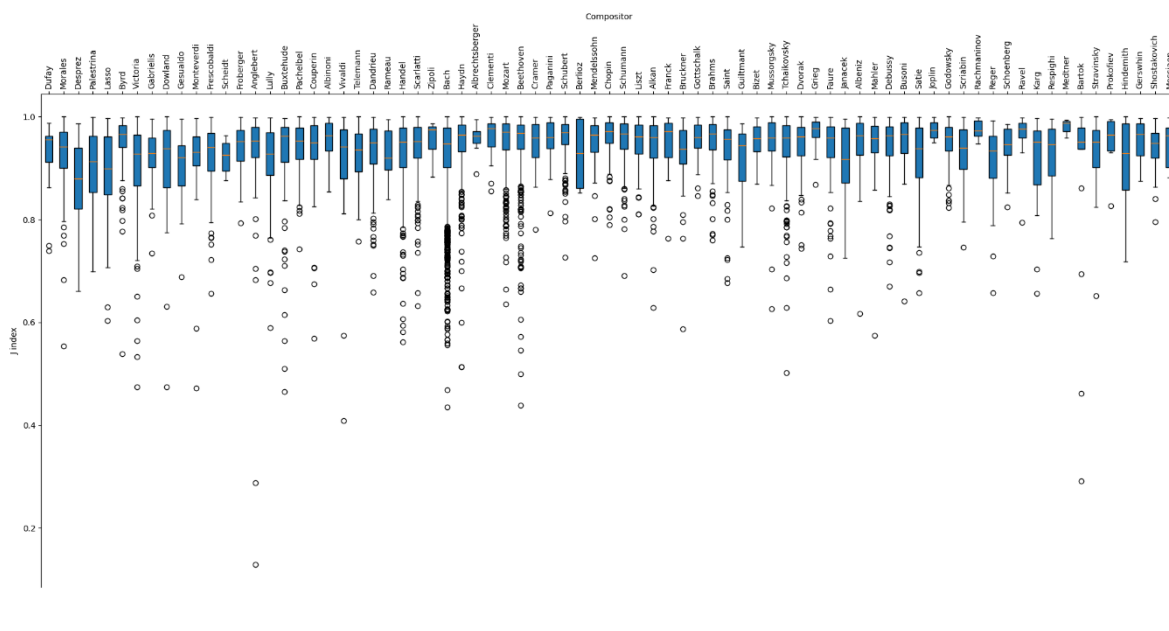


El valor mínimo de J sube rápidamente conforme N aumenta. Una longitud reducida de un segmento de datos afecta la sensibilidad en la que puede detectar posibles estructuras deterministas. Así que un solo conjunto suficientemente grande de fases aleatorias es suficiente para la comparación estadística con datos del mundo real. Los valores de J que se encuentran por debajo de esta curva mínima son fuertes indicadores de determinismo.

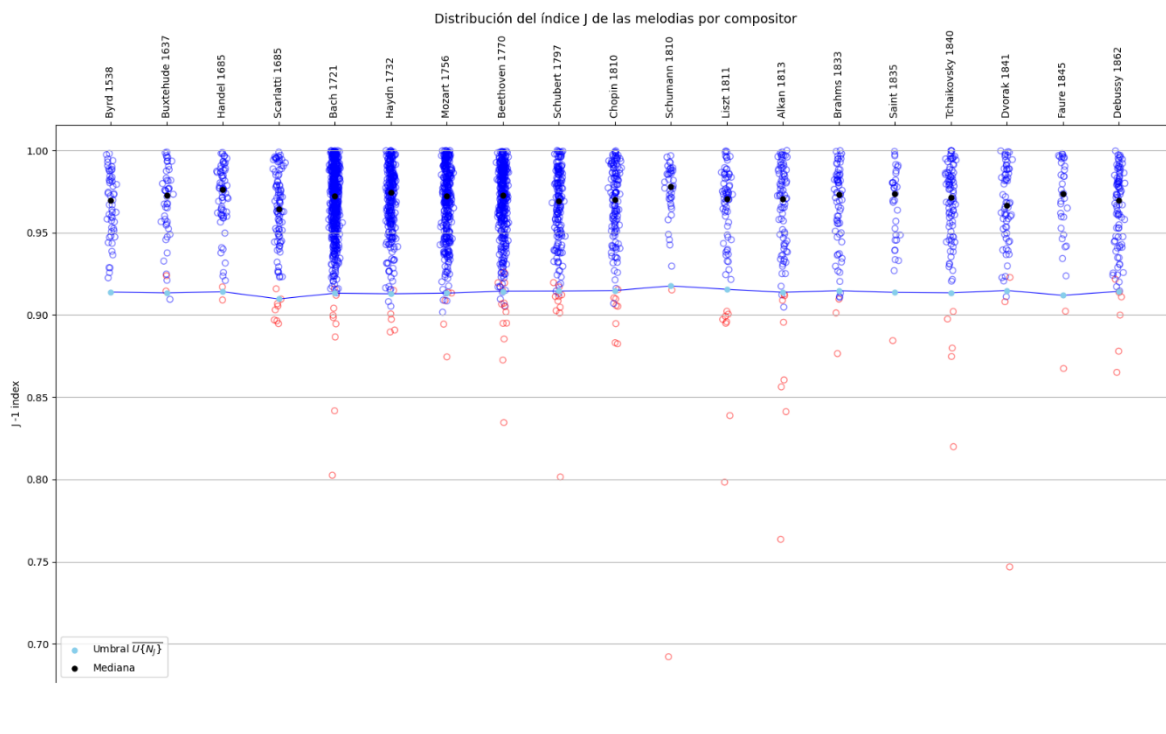
Resultados

Lo primero que se hizo fue obtener el conjunto de datos de las piezas musicales. Son archivos de texto separados por compositor. La estructura de cada uno de ellos consta de secuencias de notas representadas por números del lenguaje MIDI, los cuales representan las piezas melódicas del compositor. Cada nota es separada por renglones y cada pieza melódica es separada por un encabezado que indica el año de nacimiento y el número de piezas que hay en el archivo.

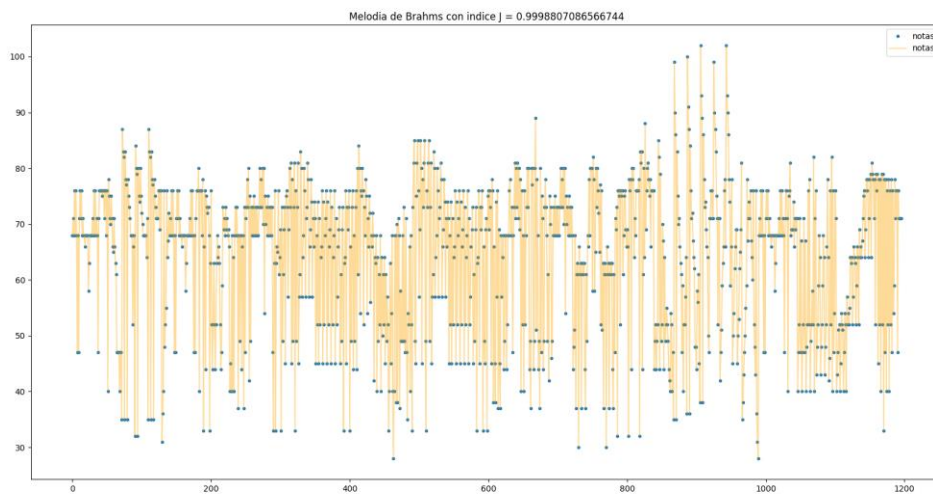
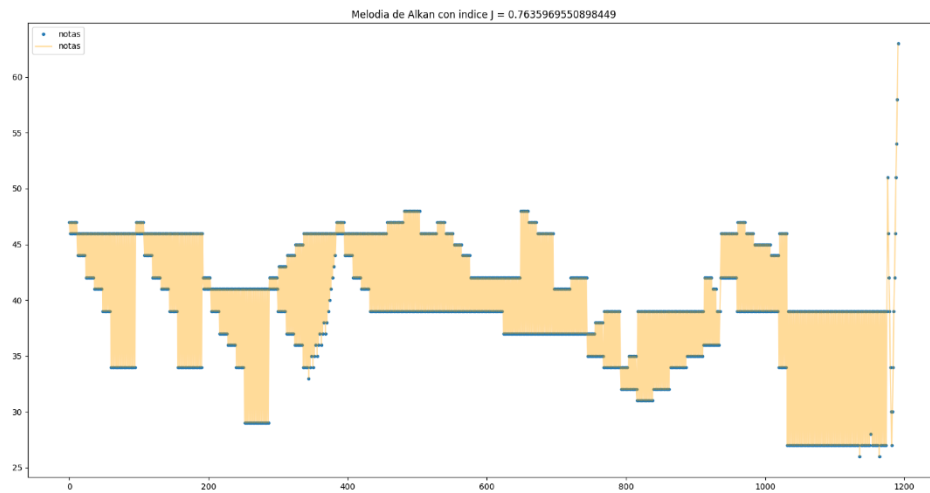
Una vez que se escribió el script de Python para guardar cada pieza melódica en listas agrupadas por un diccionario de Python, en seguida se propuso a calcular el índice J para cada una de ellas. Nótese que se tuvo cuidado en observar que no se formara el artefacto dentro de las fases de Fourier de las series de tiempo.



Hay que recordar que una de las condiciones para que se pueda calcular el índice J en una serie de tiempo, es que la cantidad de puntos en las fases de Fourier debe ser alta. El problema con este conjunto de datos es que muchas piezas melódicas no pasan de las 1000 notas, que representan 500 puntos en las fases de Fourier. Para no calcular el índice en piezas cortas de duración, se tuvo que excluirlas, lo que provoca que de los 77 compositores del conjunto de datos, solo nos quedamos con 19 de estos.



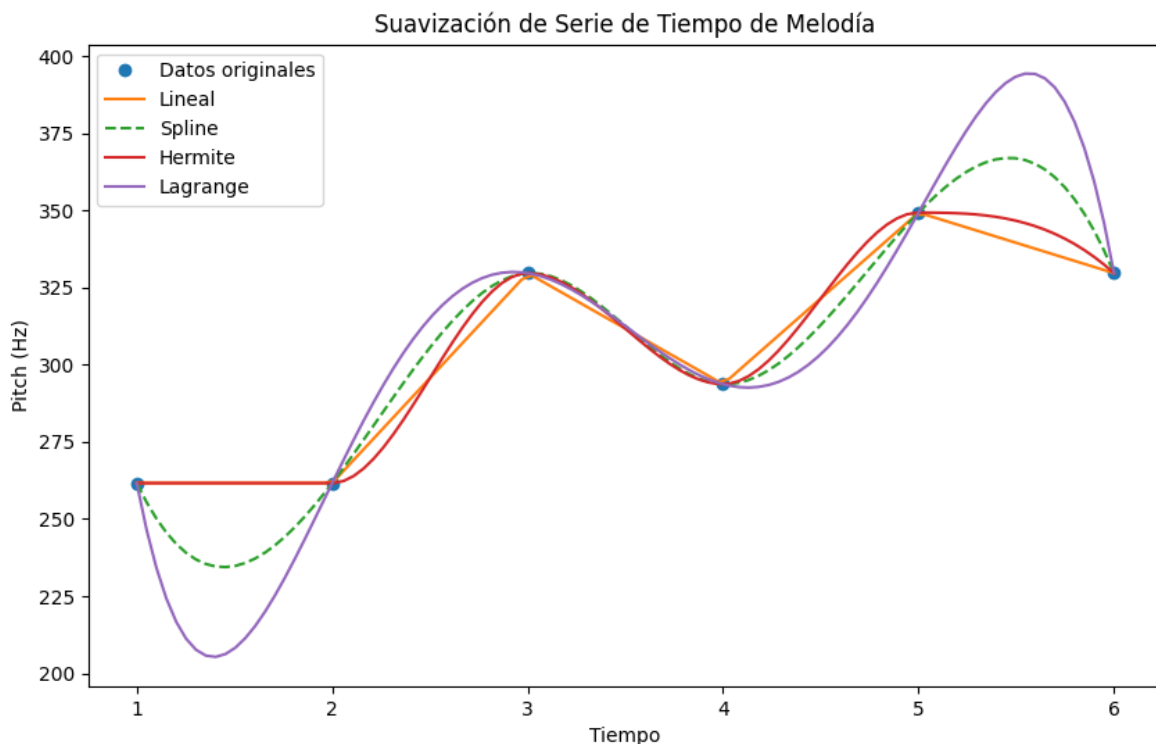
Aún así, no se distingue nada de información que pueda diferenciar las distintas corrientes de composición, o que el índice J pueda extraer componentes deterministas desde la armonía musical de las series de tiempo. Lo que sí se ha notado, es que aquellas piezas donde numéricamente se pueden apreciar una sección periódica, se calculan índices por debajo del umbral de la hipótesis nula.



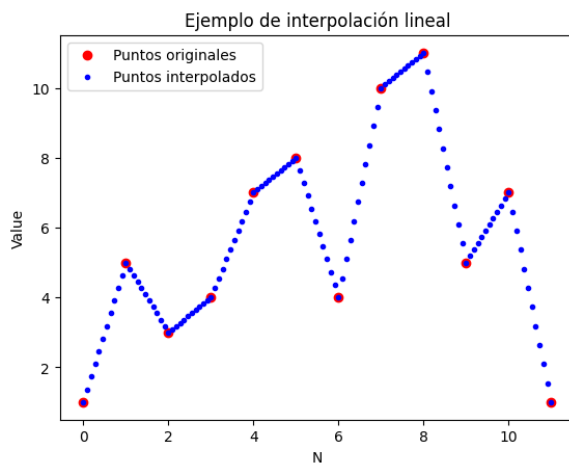
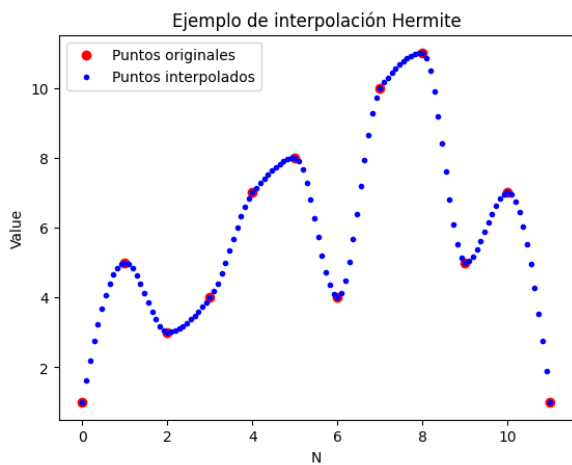
Estas son las secuencias melódicas de una pieza de Alkan y Brahms. La de Alkan se pueden apreciar secciones periódicas, a diferencia de la secuencia de Brahms donde no se puede distinguir ningún periodo. Esto puede indicar que la J , en el caso de secuencias melódicas, solo puede detectar periodicidad, y que cualquier otra secuencia, aunque se base dentro de fuertes reglas armónicas, la distinguirá como aleatoria.

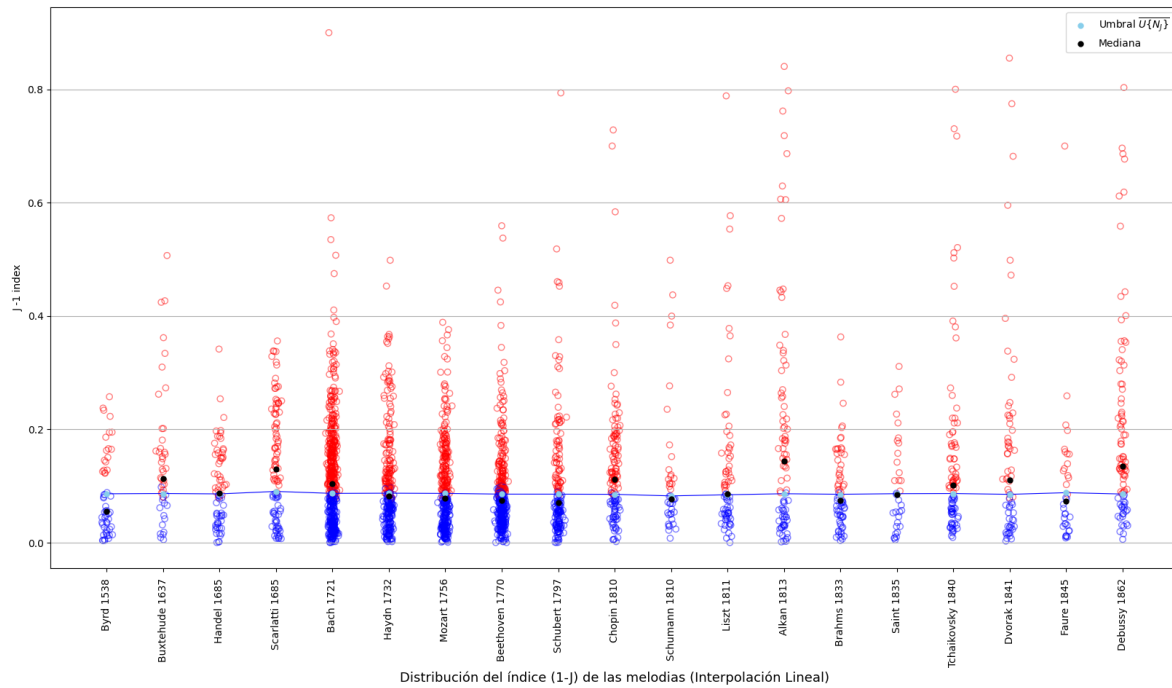
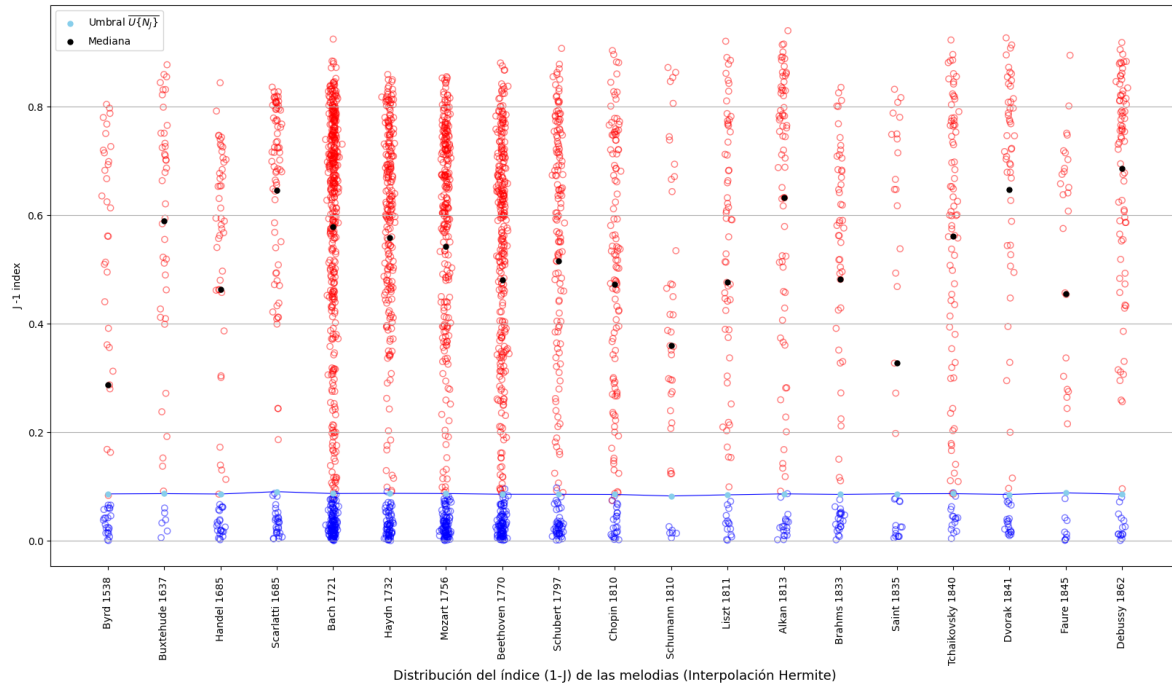
En primer momento, se pensó que se podría tratar de un problema aislado con series de tiempo discreto, donde la J no funciona en esas condiciones. Entonces lo que se hizo fue realizar una interpolación entre cada nota, con el fin de convertir la serie discreta a una

continua. La justificación musical en el uso de la interpolación, es la técnica conocida como glissando, un efecto en el que el sonido "desliza" suavemente de una nota a otra sin pausas entre ellas. Instrumentos como el trombón son capaces de hacerlo. Al deslizar la vara, el trombonista puede cambiar gradualmente la frecuencia de la nota.



Existen distintos tipos de interpolación, pero se decidió utilizar la interpolación Hermite y la interpolación lineal. La hermite conserva la información de la derivada entre los puntos, lo que genera un camino suave entre las notas. La lineal es una interpolación directa y simple.



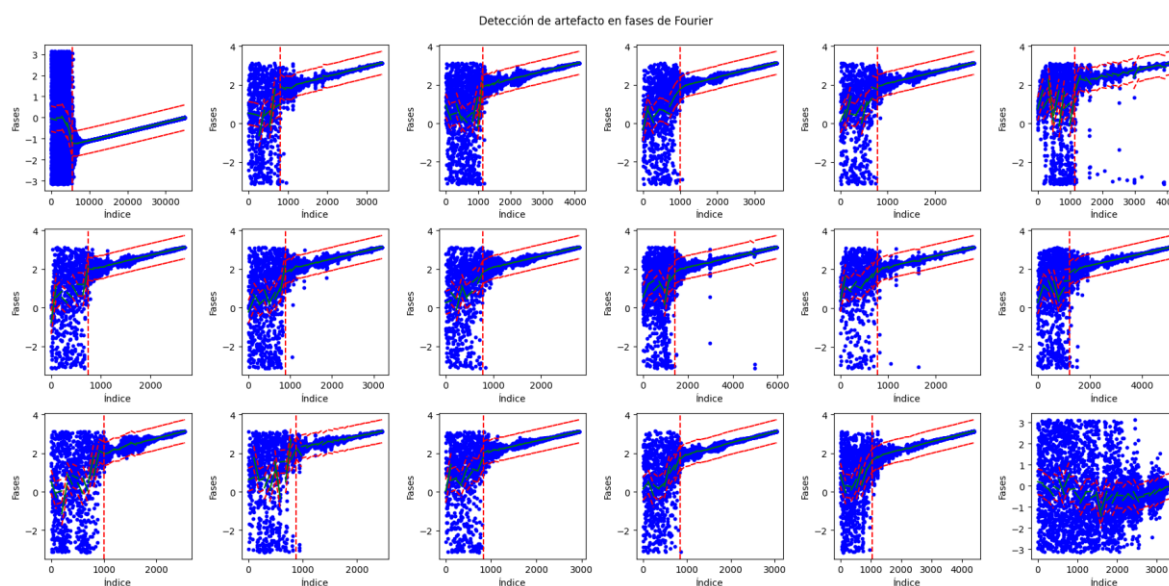


Al aplicar la interpolación a las piezas musicales, redujo los índices J por debajo del umbral para varias piezas musicales. La Hermite fue más drástica que la lineal.

Sobre las fases de Fourier

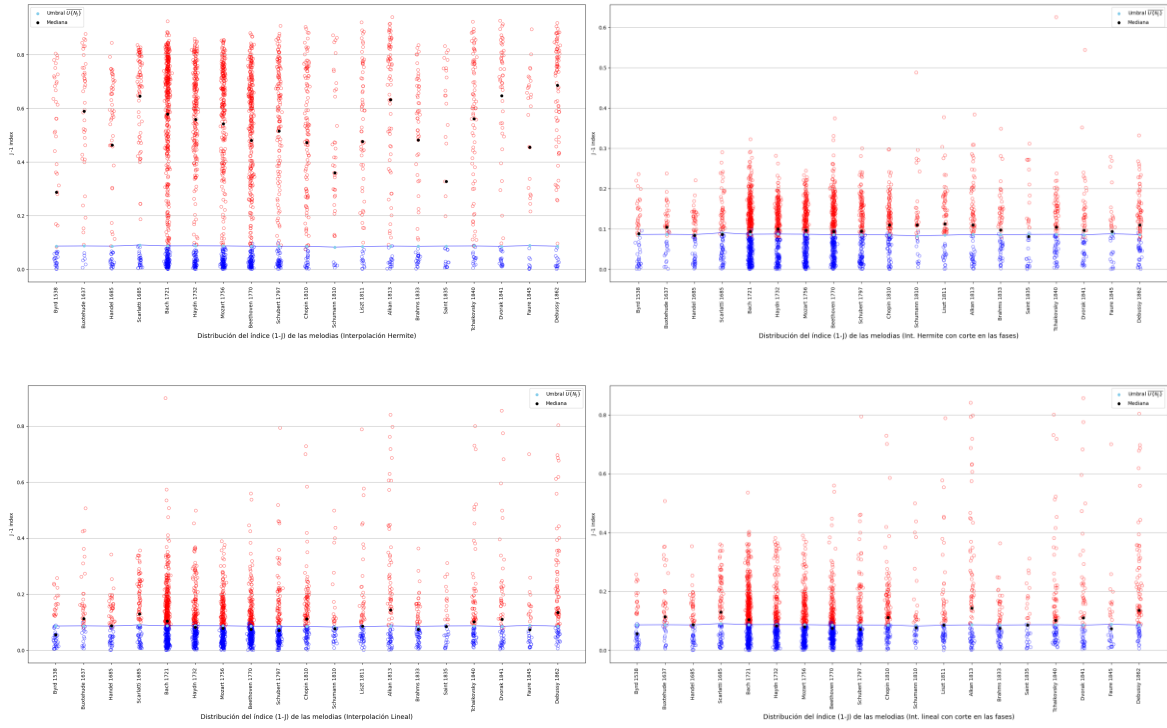
Existe el problema de que, cuando el artefacto aparece en las fases de Fourier, el índice J toma valores bajos, incluso en dinámicas aleatorias. Por ello, es necesario desarrollar un algoritmo capaz de detectar el artefacto y realizar un recorte en las fases, garantizando que el índice J no se vea afectado.

El código divide los datos en 25 ventanas de tamaño uniforme y realiza ajustes lineales en cada una. Para cada ventana, se define una región delimitada por dos líneas paralelas al ajuste lineal, desplazadas en su intersección con el eje 'y'. A continuación, se calcula el porcentaje de puntos que caen dentro de esta región con respecto al total de puntos en la ventana. Si más del 80% de los puntos están contenidos dentro de la región y este porcentaje se mantiene por encima de ese umbral en las siguientes ventanas, se identifica la presencia del artefacto.



La figura muestra las fases de Fourier correspondientes a 17 series que presentan el artefacto y una serie que no lo presenta. La línea roja vertical indica el punto de corte en la secuencia de fases, antes del inicio del artefacto. En la última gráfica, no se realizó ningún corte, ya que el artefacto no está presente.

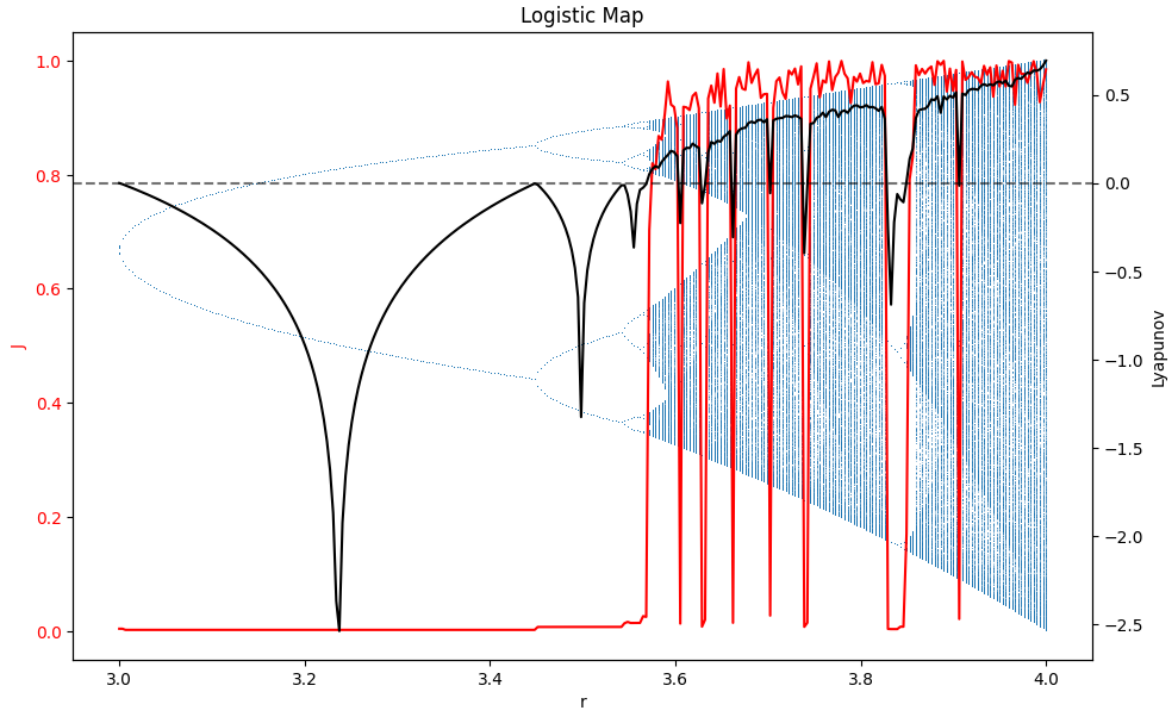
Se volvieron a calcular los índices J para las piezas de las melodías, eliminando los artefactos en las fases utilizando este algoritmo. Aún así, hay índices que no fueron afectados significativamente y otros que sí, reduciendo el número por debajo del umbral. Se desconoce la razón.



Mapecto Logístico

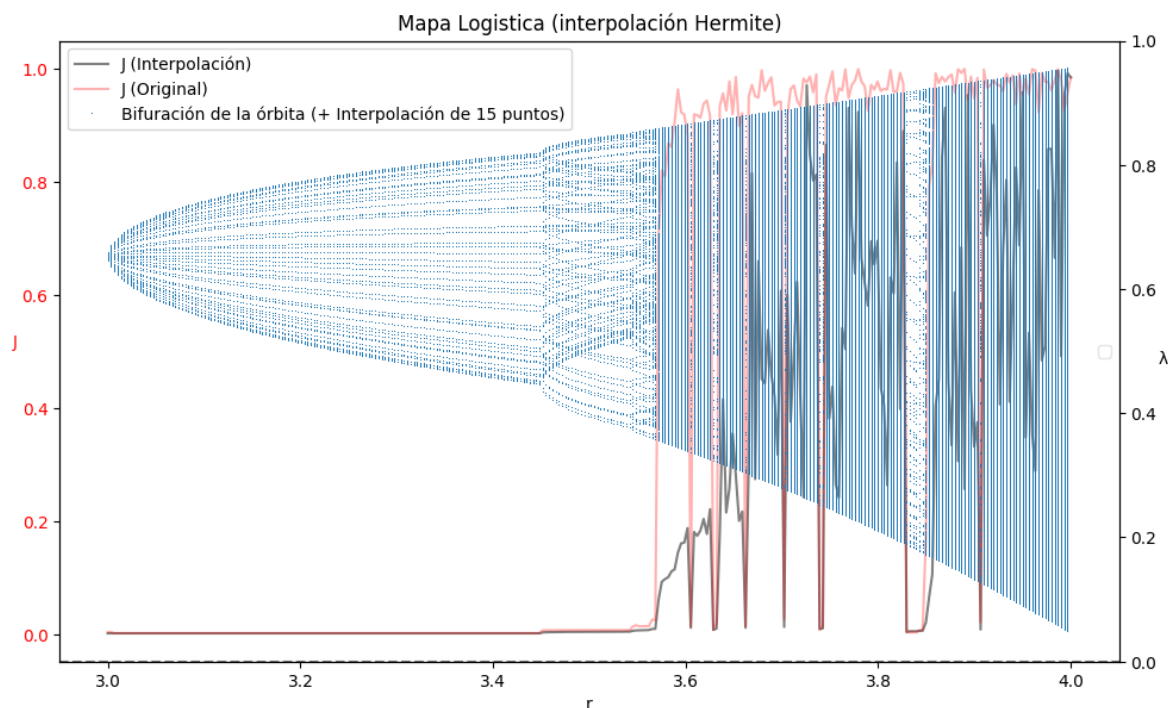
Para poder comparar el efecto de la interpolación en estas series de tiempo, se realizaron experimentos con el mapeo logístico.

$$x_{n+1} = r x_n (1 - x_n)$$



Estamos observando en azul la bifurcación de la órbita del mapa logístico en función del parámetro r , en negro tenemos el exponente de Lyapunov, donde por debajo de 0 describe una dinámica periódica y mayor que cero indica una dinámica caótica. En rojo se aprecia el índice J calculado de la órbita en función de r , el cual describe la transición entre el régimen aleatorio y regular en los mismos valores de r donde el Lyapunov describe una transición entre la dinámica caótica y periódica.

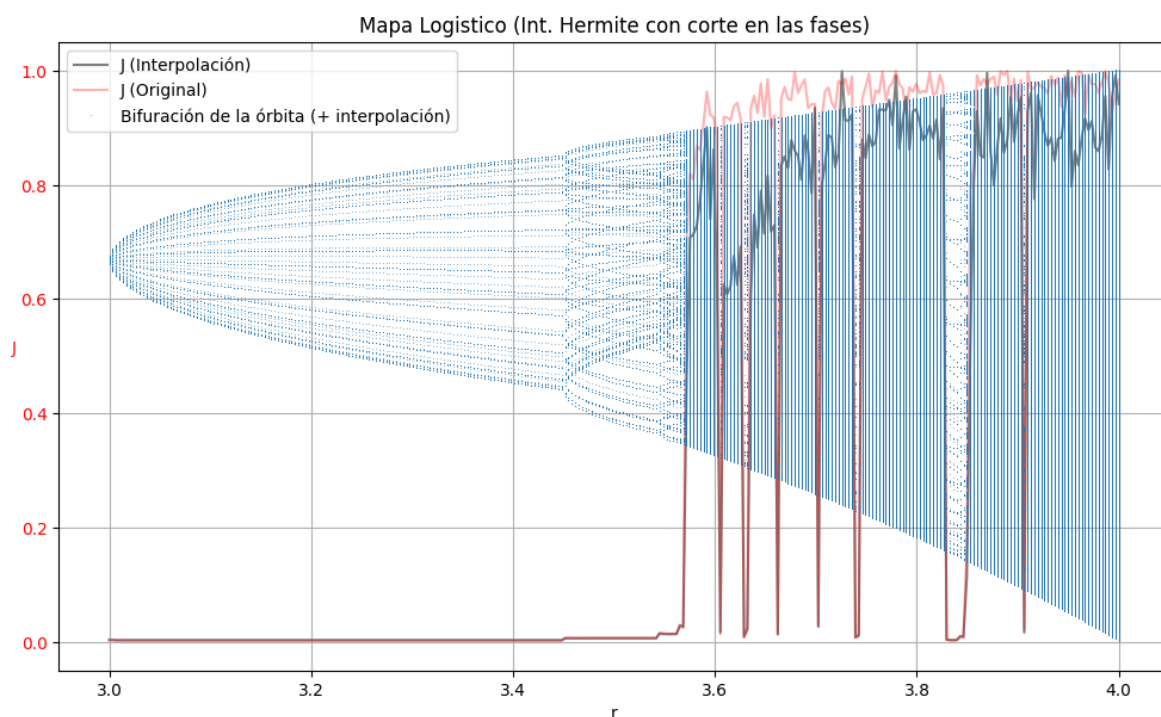
Entonces, se implementó la interpolación Hermite y lineal para cada órbita, y se comparó con la trayectoria del índice J para la órbita original.



Se presenta gráficamente la interpolación de Hermite, donde se añaden 15 puntos entre cada par de puntos originales. Esta implementación hace que el índice J pierda su capacidad para detectar dinámicas caóticas en el mapa logístico, lo que vuelve borrosa la distinción entre lo caótico y lo regular. Esto está relacionado con el efecto observado al aplicar la interpolación de Hermite en piezas musicales, donde los valores de J caían por debajo del umbral tras la implementación.

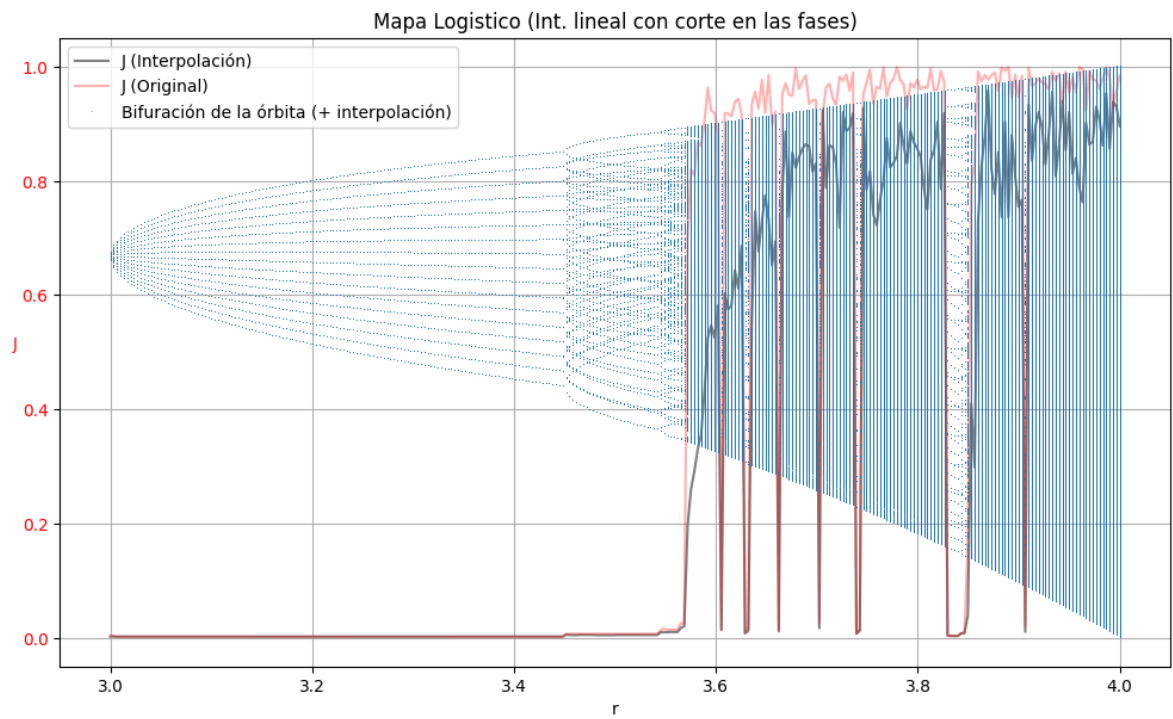
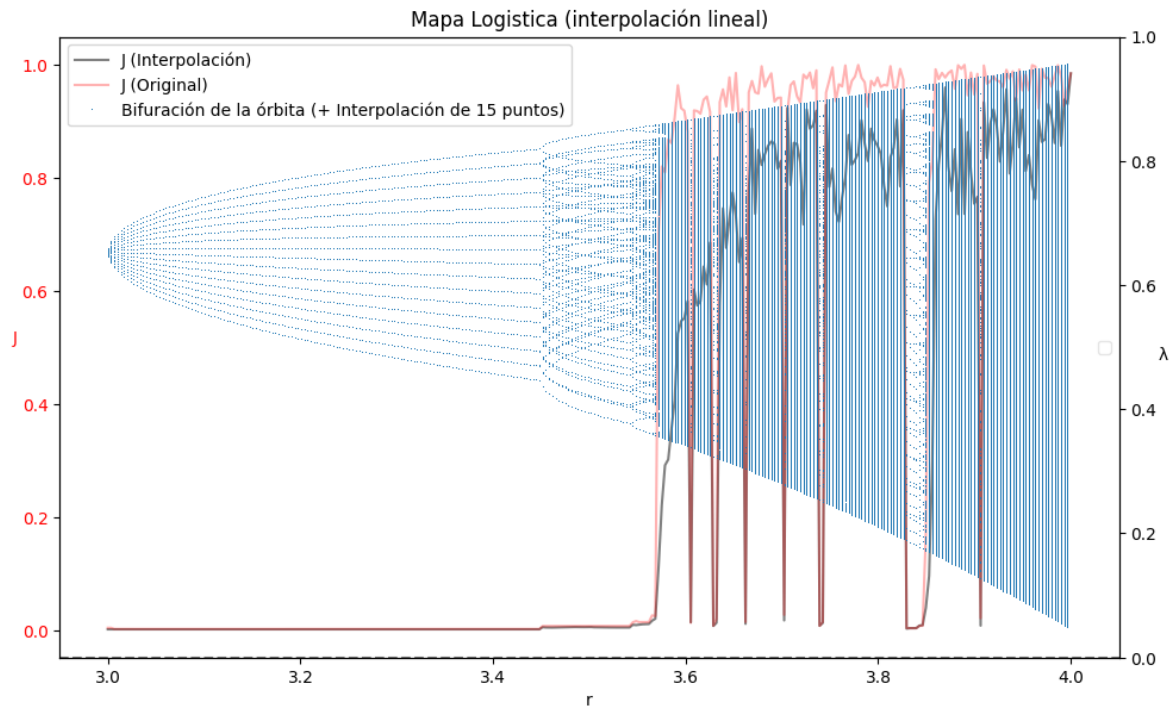
Sin embargo, al incorporar el algoritmo de detección y eliminación del artefacto en las fases de Fourier al cálculo del índice J , los resultados se vuelven más estables. Las regiones periódicas no se vieron afectadas, confirmando la ausencia de artefactos en dichas áreas. En contraste, en las regiones caóticas, donde sí había artefactos, la eliminación de estos permitió observar la reducción del índice J debido únicamente al componente determinista añadido por la interpolación, y no por el efecto de artefactos. La distinción entre lo caótico y lo regular

se vuelve más clara tras la remoción de los artefactos.



Se aplicó el mismo procedimiento utilizando interpolación lineal, y se observó que la diferencia entre el índice J calculado con el corte en las fases de Fourier y el índice calculado sin el corte es menos pronunciada en comparación con la interpolación de Hermite. Esto se debe a que la interpolación lineal generó significativamente menos casos de artefactos en las fases que la interpolación de Hermite.

Con esto se puede concluir que la incorporación de interpolaciones basadas en componentes deterministas en dinámicas completamente aleatorias impacta el cálculo del índice J , haciéndolo capaz de detectar dichas interpolaciones y reduciendo su valor por debajo del umbral correspondiente a la hipótesis nula.



Si el índice J es afectado por implementación de un componente determinista en una dinámica aleatoria, entonces, ¿Será afectada por la implementación de un componente aleatorio en una dinámica regular?

Para esto, se quiso calcular el índice J para el mapa logístico, pero con ruido añadido. Ahora bien, hay dos formas de añadir ruido a la órbita, una de ellas, llamada ruido multiplicativo, hace que la varianza del ruido sea dependiente de la intensidad del atractor de la órbita, y la otra no, la que se llamará ruido aditivo, mantiene a la varianza del ruido η_t igual.

$$X_{t+1} = X_t r (1 - X_t)$$

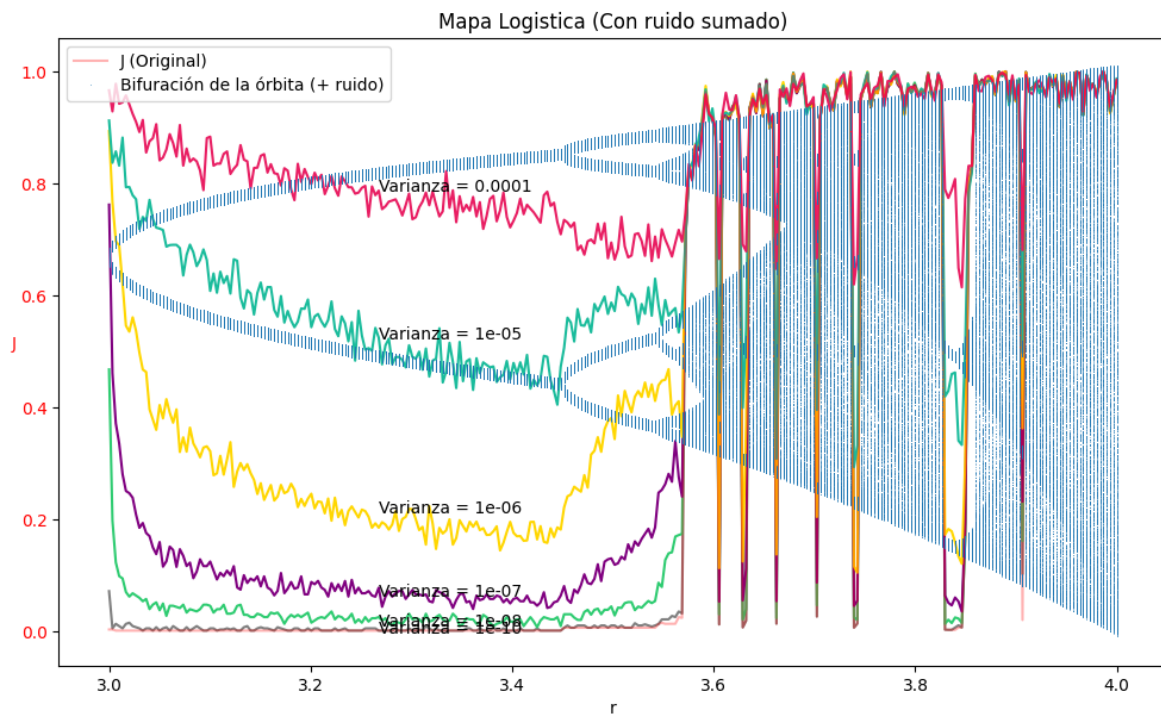
Ruido aditivo

$$X'_{t+1} = X_{t+1} + \eta_t$$

$$X'_{t+1} = X_{t+1} + \eta_t$$

$$X'_{t+2} = X_{t+2} + \eta_{t+1}$$

$$X'_{t+3} = X_{t+3} + \eta_{t+2}$$

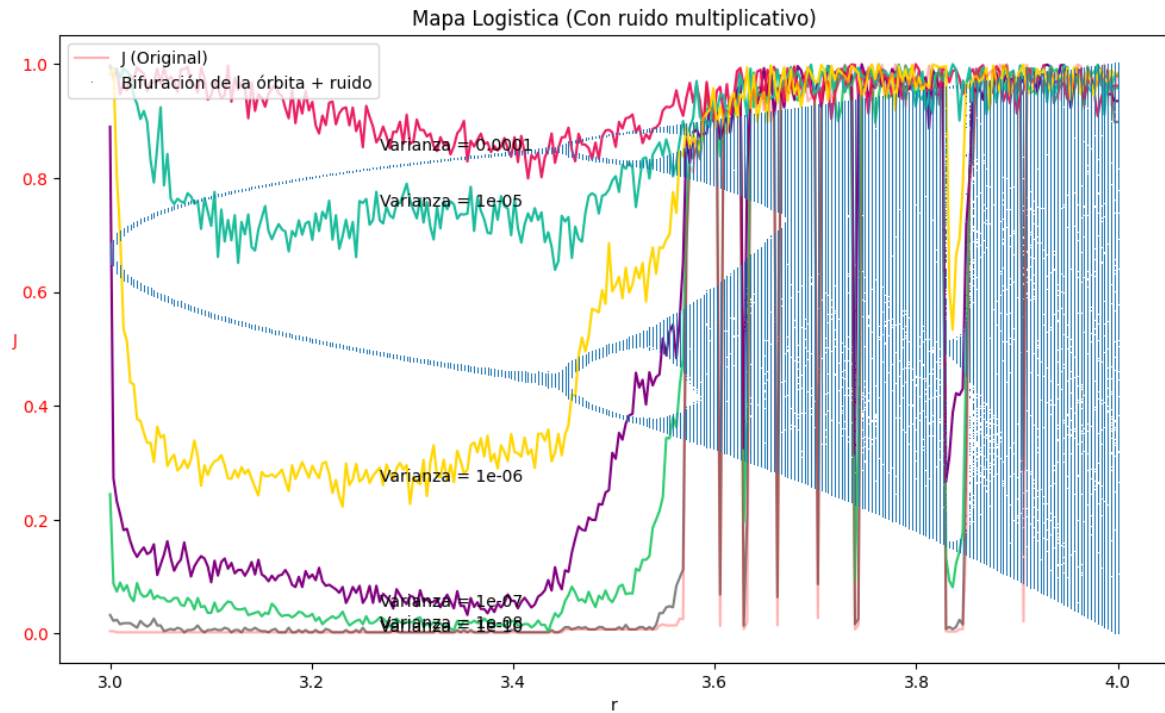


Ruido multiplicativo

$$X'_{t+1} = X_{t+1} + \eta_t$$

$$X'_{t+2} = X'_{t+1} + \eta_{t+1}$$

$$X'_{t+3} = X'_{t+2} + \eta_{t+2}$$



El grosor en donde normal solo se ve un punto, representa la varianza del ruido, y en ejemplo de ruido multiplicativo, este grosor varía dependiendo de la intensidad de los atractores en la dinámica. En el ejemplo del ruido aditivo, este grosor se mantiene igual.

Interpolación estocástica

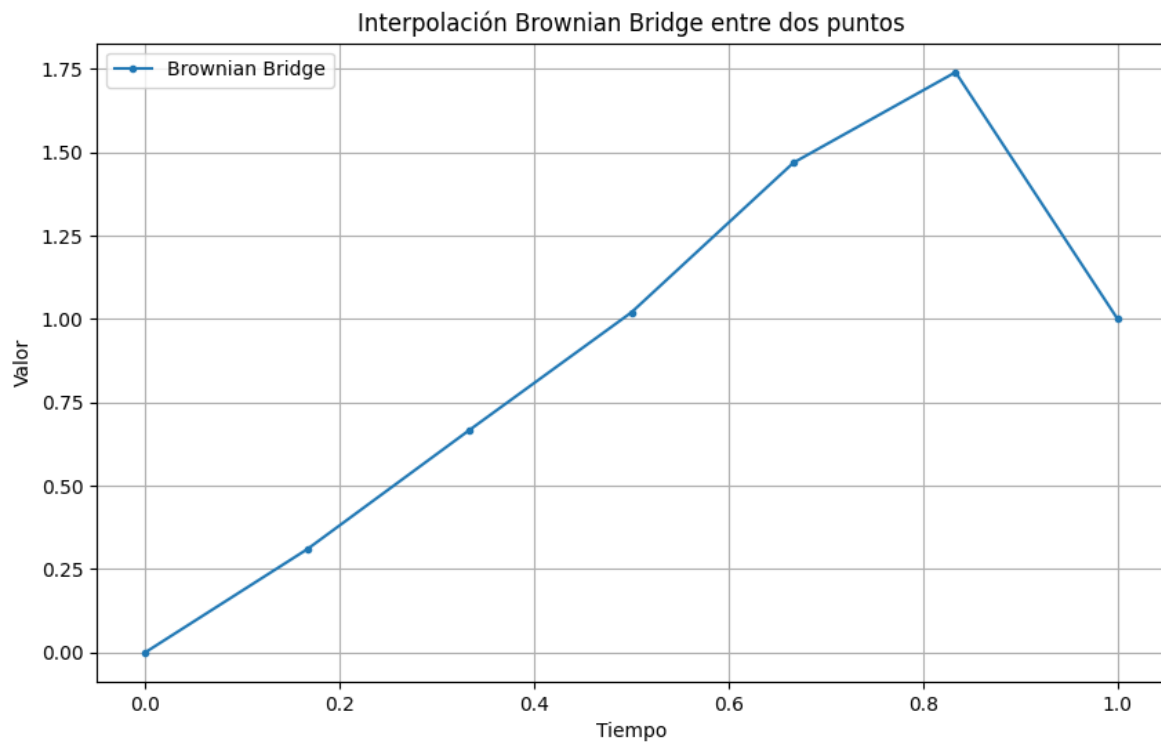
La interpolación estocástica es un método de interpolación entre dos puntos mediante un proceso estocástico, manteniendo las propiedades de un movimiento browniano (o proceso de Wiener).

$$X(T) = X(0) + \frac{t}{T}(X(T) - X(0)) + W(t) - \frac{t}{T}W(T)$$

Donde:

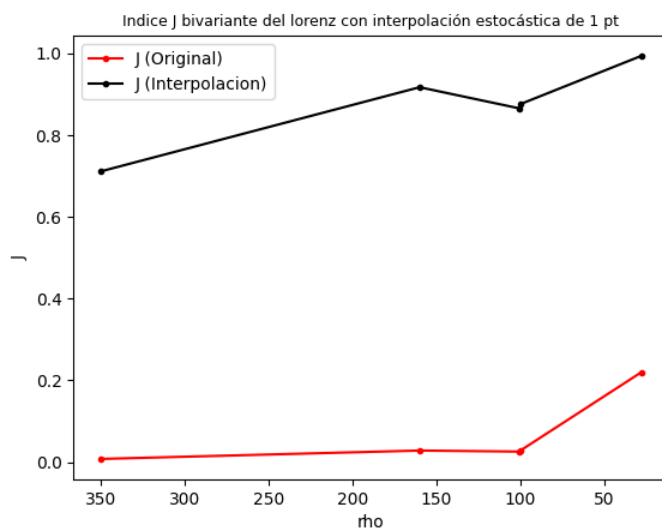
- $W(t)$ es el proceso de Wiener
- t es el tiempo inicial
- T el tiempo final
- $X(0)$ es el valor inicial

- $X(T)$ el valor final

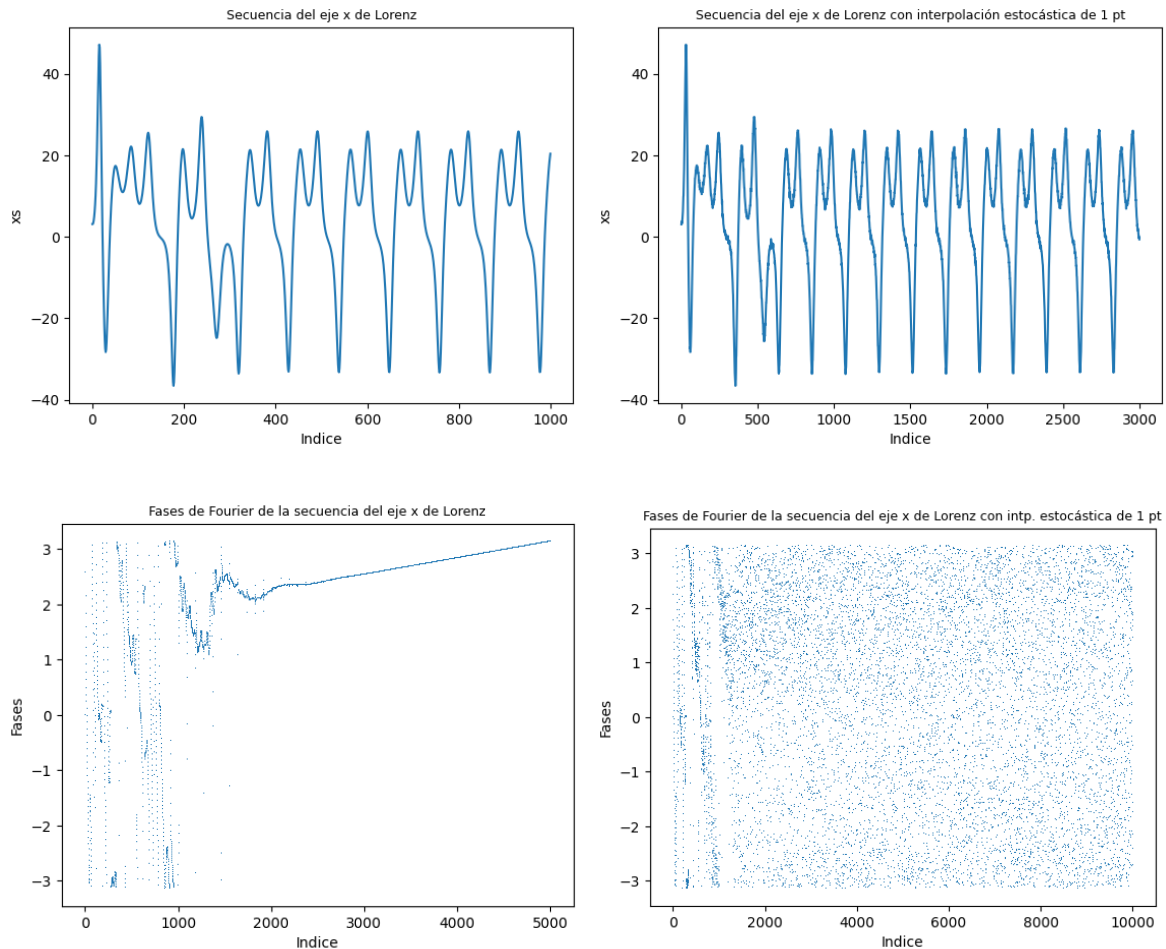


Como otra propuesta para observar el efecto del índice J ante componentes aleatorios o estocásticos, se decidió realizar experimentos de la implementación de la interpolación estocástica en sistemas dinámicos continuos: sistema de Lorenz y Rossler

Tras interpolar tres puntos entre cada punto original del sistema de Lorenz y calcular el índice J bivalente entre los ejes 'x' y 'y', se observó un incremento en el valor del índice J .



Al revisar las fases de Fourier de la secuencia del eje 'x' y las fases de la secuencia con la interpolación estocástica cuando $\rho=100.5$, se puede observar la razón del aumento del índice J. En la secuencia original, el índice J univariante es de 0.06, mientras que el índice J univariante en la secuencia con interpolación, es de 0.917.



References

- [1] González-Espinoza, A., Larralde, H., Martínez-Mekler, G., & Müller, M. (2017). Multiple scaling behaviour and nonlinear traits in music scores. Royal Society Open Science, 4(12), 171282. <https://doi.org/10.1098/rsos.171282>
- [2] González-Espinoza, A., Martín ez-Mekler, G., & Lacasa, L. (2020). Arrow of time across five centuries of classical music. Physical Review Research, 2(3). <https://doi.org/10.1103/physrevresearch.2.033166>

[3] A detailed description of the algorithm and its application to physiologic signals can be found in: Peng C-K, Havlin S, Stanley HE, Goldberger AL. [Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series](#). *Chaos* 1995;**5**:82-87.

[4] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* **101**(23):e215-e220 [Circulation Electronic Pages; <http://circ.ahajournals.org/content/101/23/e215.full>]