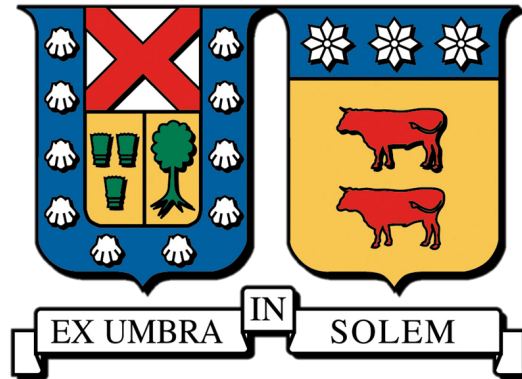


UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

DEPARTAMENTO DE INGENIERÍA MECÁNICA



**APLICACIÓN DE TÉCNICAS DE TRANSPORTE ÓPTIMO
EN PROCESO DE FORMACIÓN DE IMÁGENES DE ULTRASONIDO**

David Rodolfo Canela Espinosa

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL MECÁNICO

Guía: Dr. Joaquín Múra
Co-referentes: Dr. Felipe Galarce

JUNIO-2026



CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): ☒ Memoria o trabajo de título ☐ Tesis de Postgrado

Título del trabajo: Aplicación de técnicas de transporte óptimo en proceso de formación de imágenes de ultrasonido

Nombre del candidato(a): David Rodolfo Canela Espinosa

Carrera / Grado: Ingeniería Civil Mecánica

Campus: San Joaquín

Departamento: Departamento Ing. Mecánica Campus San Joaquín

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Joaquín Mura Mardones, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

☒ El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

☐ El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por (marcar una opción):

☐ 6 meses ☐ 12 meses ☐ 2 años ☐ 3 años ☐ 5 años ☐ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 19/08/2026

Firma: [Firma]

Estudiante o Candidato(a): David Rodolfo Canela Espinosa

Fecha: 18/08/2026

Firma: [Firma]

Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.

Resumen

El presente documento busca evaluar la validez de la aplicación de interpolaciones que aprovechan la propiedad de conservar estructuras geométricas, propia de la resolución de problemas de transporte óptimo, en señales ultrasónicas pertenecientes a las distintas etapas de la construcción de imágenes modo B. La principal limitación de estas interpolaciones es tener que trabajar sobre distribuciones de probabilidad, lo cual obliga a realizar un proceso de normalización y posterior reescalamiento. La etapa sobre la cual se realizaron interpolaciones que dio mejores resultados fue sobre matrices modo B, donde se aumentó la resolución temporal mediante interpolación de frames para una adquisición de datos ultrasónicos.

Agradecimientos

Me gustaría agradecer a mi familia, especialmente a mi mamá Dora y mi abuela Gigi por aguantarme, prestarme apoyo y amor incondicional. Agradecer a todas las personas que mi trayecto a lo largo de la universidad me permitio conocer. Mis mejores amigos Tamara y Gonzalo por alentarme de manera constante. Y a mi gato Juancho por siempre estar a mi lado mientras escribí este documento.

Agradecer también a todos los integrantes del Laboratorio de Ultrasonido Cuantitativo de la USM y al proyecto FONDECYT N.º 1230864.

Tabla de Contenidos

1	Introducción	1
2	Objetivos y Alcances	2
3	Marco Teórico	4
3.1	Ondas Mecánicas	4
3.1.1	Tipos de ondas mecánicas	4
3.1.2	Parámetros	6
3.1.3	Interacción con medios de propagación	7
3.1.4	Representaciones Tiempo-Frecuenciales	8
3.2	Ultrasonido	11
3.2.1	Transductores	11
3.2.2	Transmisión y recepción	12
3.2.3	Construcción imagen B-mode	12
3.3	Transporte óptimo	15
3.3.1	Formulación de Monge	16
3.3.2	Formulación de Kantorovich	17
3.3.3	Distancias y barycentros de Wasserstein	21
3.3.4	Transporte óptimo unidimensional	23
3.3.5	Formulación de Kantorovich regularizada entrópicamente	25
3.3.6	Transporte óptimo desbalanceado	28
4	Metodología	31
4.1	Interpolación sobre señal RF	31
4.1.1	Interpolación entre espectrogramas	32
4.1.2	Interpolación distancias TL^p	36
4.1.3	Evaluación de interpolaciones OT sobre señales RF	38
4.2	Interpolación sobre envolventes post-beamforming	38

4.3	Interpolación sobre matrices B-mode	40
4.4	Evaluación de interpolaciones	41
4.4.1	Interpolación sobre señales RF	42
4.4.2	Interpolación sobre envolventes post-beamforming	43
4.4.3	Interpolación sobre matrices B-mode	43
5	Resultados	44
5.1	Interpolación sobre señal RF	44
5.1.1	Interpolación entre espectrogramas	45
5.1.2	Interpolación con distancias TL^p	47
5.1.3	Análisis de resultados	48
5.2	Interpolación sobre envolventes post-beamforming	50
5.2.1	Análisis de resultados	53
5.3	Interpolación sobre matrices B-mode	55
5.3.1	Análisis de resultados	56
6	Conclusiones y Recomendaciones	59
	Referencias	61
A	Códigos metodología	I

Lista de Figuras

3.1	Onda Longitudinal	5
3.2	Onda Transversal	5
3.3	Amplitud y longitud de onda	6
3.4	Interacción onda-interfaz	7
3.5	Difracción de onda	8
3.6	Transductor lineal, convexo y de matriz en fase	12
3.7	Ilustración del mapa de transporte de Monge, ν es el pushforward de μ según T . .	16
3.8	Plan de transporte π entre distribuciones continuas μ y ν	18
3.9	Ilustración del espacio métrico que W_p define sobre el espacio de medidas de probabilidad de momento de orden p finito	22
3.10	Comparación de interpolación inducida por plan de transporte óptimo π con interpolación lineal para un mismo par de distribuciones de probabilidad	23
4.1	Señales 1 y 2 para ejemplificar el metodo de interpolación OT sobre espectrogramas	33
4.2	Espectrogramas de magnitud de las señales 1 y 2.	34
4.3	Distribuciones de probabilidad correspondientes a la energia espectral de s_1 y s_2 para $t = 0$	35
4.4	Distribución de probabilidad interpolada con factor $\alpha = 0.5$ entre columnas frecuenciales de espectrogramas de s_1 y s_2 para $t = 0$	35
4.5	Espectro de magnitud interpolado mediante tecnicas OT entre las señales s_1 y s_2 .	36
4.6	Señal interpolada entre s_1 y s_2 mediante OT en espectros de magnitud	36
4.7	Elementos de transductor para evaluación de interpolación sobre señales RF. En azul se encuentran los elementos sin alterar, mientras que en rojo los elementos sobre los cuales se sobreescribe una interpolación	42
5.1	Imagen B-mode construida a partir de datos RF no intervenidos	45
5.2	Imagen B-mode construida a partir de datos RF con señales RF interpoladas linealmente	45

5.3	Imagen B-mode construida a partir de datos RF con señales RF interpoladas por OT regularizado	46
5.4	Imagen B-mode construida a partir de datos RF con señales RF interpoladas por OT con distancias TL^p	48
5.5	RMSE de señales RF interpoladas con respecto a señales RF originales	49
5.6	Distancia de Wasserstein orden 2 entre señales RF interpoladas y señales RF originales	49
5.7	Imagen B-mode original y con columnas impares interpoladas linealmente	50
5.8	Imagen B-mode con columnas impares reemplazadas por OT unidimensional	51
5.9	Imagen B-mode con columnas impares reemplazadas por UOT $\rho = 0.1$	52
5.10	Imagen B-mode con columnas impares reemplazadas por UOT $\rho = 0.99$	53
5.11	RMSE entre columnas interpoladas y originales	54
5.12	Distancia de Wasserstein orden 2 entre columnas interpoladas y originales	54
5.13	RMSE entre frames interpoladas y originales	57
5.14	PSNR entre frames interpoladas y originales	57
5.15	SSIM entre frames interpoladas y originales	57

Capítulo 1

Introducción

Las imágenes ecográficas constituyen una herramienta fundamental en el diagnóstico y seguimiento clínico debido a su carácter no invasivo, su bajo costo relativo y su capacidad de generar información en tiempo real. Similar a la ecolocalización que utilizan submarinos o murciélagos para ubicarse en el espacio, la formación de imágenes ultrasónicas modo B se basa en la transmisión y posterior recepción de ondas acústicas mediante un arreglo de elementos piezo-eléctricos ubicados en un transductor. Estas señales recibidas pasan por un proceso de formación de haz Delay-and-Sum y una compresión logarítmica para llegar a una imagen, la cual representa de forma espacial la intensidad de los ecos recibidos desde distintas regiones del tejido estudiado.

Sin embargo, la calidad de la imagen se encuentra condicionada por limitaciones físicas y tecnológicas. Principalmente el número finito de elementos piezoeléctricos disponibles en el transductor y la respectiva separación espacial que existe entre cada uno de ellos. Estas restricciones limitan la resolución lateral y temporal de la imagen, mermando la capacidad representar estructuras pequeñas o perjudicar la continuidad de la información adquirida. Una alternativa para enfrentar estas limitaciones consiste en estimar señales, columnas o frames intermedios a partir de los datos originales disponibles, aumentando virtualmente la capacidad de muestro del hardware del sistema.

Capítulo 2

Objetivos y Alcances

La formación de imágenes ecográficas de alta resolución se encuentra limitada por diversos factores físicos y tecnológicos, entre los que destaca el número finito de elementos piezoeléctricos que conforman el transductor. Esta limitación restringe el muestreo espacial de las ondas ultrasónicas y, en consecuencia, la resolución lateral y temporal que puede alcanzarse durante la adquisición, sin que sea posible aumentar indefinidamente el número de elementos debido a restricciones de costo, tamaño y complejidad electrónica. Frente a este escenario, las técnicas de interpolación surgen como una alternativa computacional para estimar información intermedia a partir de los datos adquiridos, permitiendo incrementar la densidad de muestreo sin modificar el hardware del sistema. En este contexto, el presente trabajo tiene como propósito estudiar el transporte óptimo como una herramienta para la interpolación de datos ultrasónicos, analizando tanto sus fundamentos teóricos como sus posibilidades de aplicación práctica. Para ello, se realiza una revisión bibliográfica de los principales conceptos asociados al transporte óptimo y a sus métodos numéricos de resolución, con especial énfasis en el algoritmo de Sinkhorn. Posteriormente, estas herramientas se aplican a la interpolación de señales RF e imágenes en modo B, evaluando su capacidad para preservar la información geométrica de las estructuras y su potencial para contribuir al aumento de la resolución de las imágenes ecográficas.

—

Capítulo 3

Marco Teórico

3.1 Ondas Mecánicas

Las ondas mecánicas son un tipo de onda, las cuales se definen como perturbaciones que se propagan a través de un medio material mediante la interacción de las partículas que la componen. Estas perturbaciones tienen un carácter cíclico producto de la elasticidad del medio, que actúa como una fuerza que obliga a las partículas a volver a su posición de equilibrio, y la inercia, la cual produce que las partículas sobrepasen su posición de equilibrio al intentar volver a ella. Además permite transmitir el movimiento a partículas vecinas, propagando la onda mecánica a través del medio.

3.1.1 Tipos de ondas mecánicas

Las ondas mecánicas se pueden subdividir en 2 grandes grupos según la dirección de propagación y la dirección de oscilación de las partículas con respecto a la dirección de propagación.

Ondas longitudinales

Las ondas longitudinales son aquellas en las que la dirección de propagación coincide con la dirección de oscilación de las partículas, provocando esfuerzos de compresión y rarefacción en el medio. Es por esto que la velocidad de propagación depende de la densidad y la constante de elasticidad.

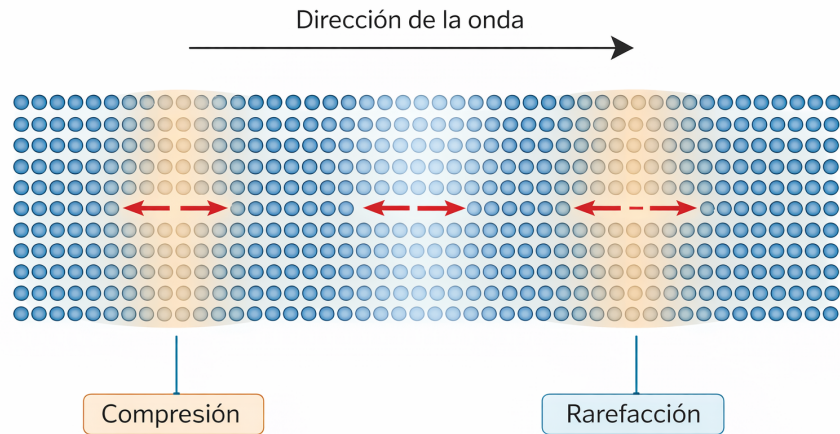


Figura 3.1: Onda Longitudinal

Ondas transversales

Las ondas transversales son aquellas en las que la dirección de propagación es perpendicular a la dirección de oscilación de las partículas, produciendo esfuerzos cortantes entre las partículas. Los gases y líquidos son particularmente incapaces de transmitir este tipo de onda dada su nula resistencia ante esfuerzos cortantes

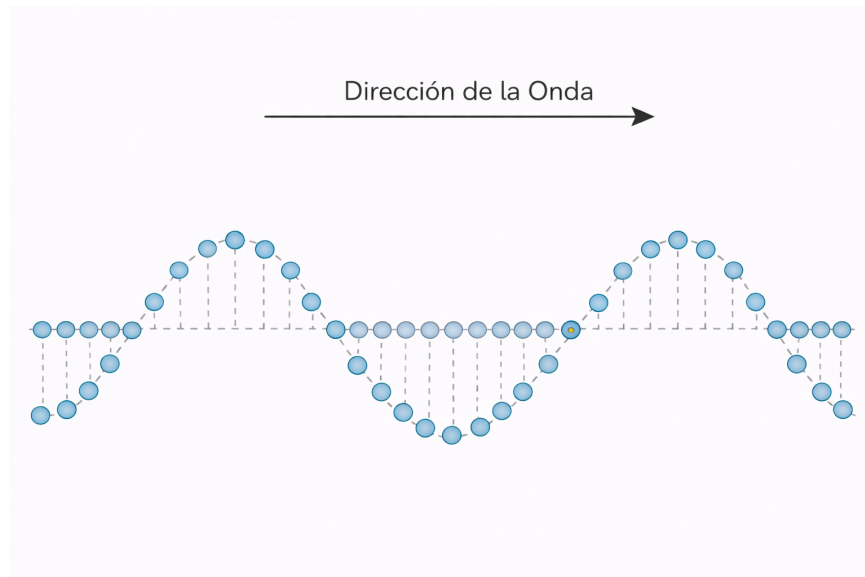


Figura 3.2: Onda Transversal

3.1.2 Parámetros

Parámetros de ondas

Los parámetros principales que permiten caracterizar una onda son los siguientes:

Amplitud: Distancia entre el punto más elongado de la onda y su punto de equilibrio.

Longitud de onda: Distancia que separa dos puntos iguales entre dos ciclos consecutivos.

Periodo: Tiempo que demora la onda en completar un ciclo.

Frecuencia: Recíproco multiplicativo del periodo. Corresponde a la cantidad de ciclos completados por unidad de tiempo.

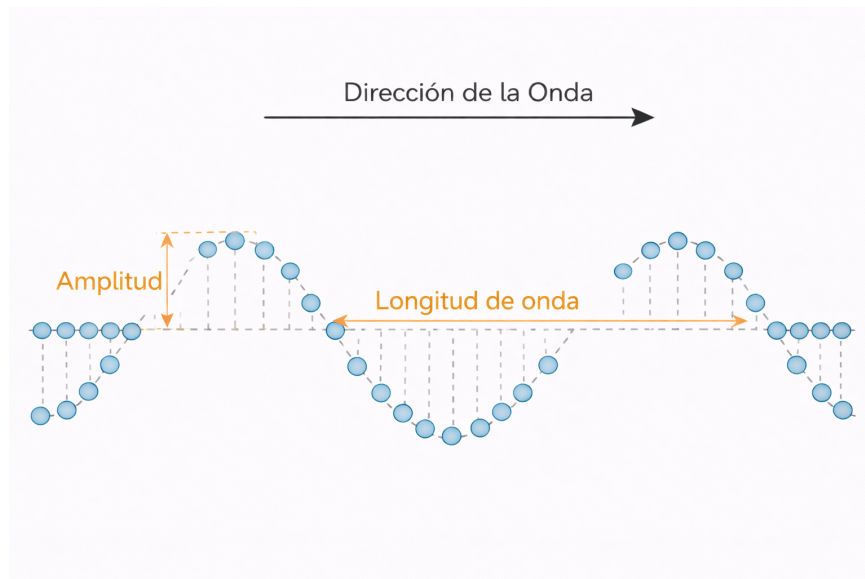


Figura 3.3: Amplitud y longitud de onda

Parámetros del medio

Velocidad sonora/acústica: Velocidad a la cual el frente de onda se propaga a través del medio material. Esta velocidad de propagación está estrechamente relacionada con las propiedades mecánicas del medio, como su densidad o sus constantes elásticas, y se puede calcular mediante la siguiente ecuación:

$$c = \lambda f \quad (3.1)$$

Velocidad instantánea: Corresponde a la velocidad propia de la partícula asociada al movimiento oscilatorio de la onda.

Presión acústica: Corresponde a la variación que experimenta la presión entre partículas, producto de la propagación de una onda a través del medio que estas componen. La amplitud de presión acústica indica la máxima variación de la presión con respecto a su valor de equilibrio.

Impedancia acústica: Este parametro permite cuantizar la oposición que presenta el medio de propagación a las oscilaciones que la onda mecánica impone sobre el. Esta puede ser expresada por el producto entre su densidad ρ y su velocidad sonora c

$$Z = \rho c \quad (3.2)$$

3.1.3 Interacción con medios de propagación

Las ondas mecánicas al propagarse a traves de diversos medios estas se enfrentan a distintos fenomenos que alteran sus características. Estos fenomenos usualmente se dan en las interfaces entre dos medios que posean una impedancia acústica distinta y por lo general obedecen las mismas leyes que rigen sobre luz y su interacción con el universo.

Transmisión, Reflexión y Refracción

Cuando una onda que se enfrenta a una interfaz entre dos medios con distinta velocidad sonora una fracción de su energia se ve reflejada en una onda que sigue propagandose a traves del medio original, mientras que la fracción energética restante atraviesa la frontera y se va en una onda de transmisión en el nuevo medio de propagación. Si la onda choca de forma oblicua con la interfaz, la onda transmitida se propagará con un angulo distinto al incidente, llamado angulo de refracción, el cual obedecerá la ley de Snell.

$$\frac{\sin(\theta_i)}{c_i} = \frac{\sin(\theta_t)}{c_t} \quad (3.3)$$

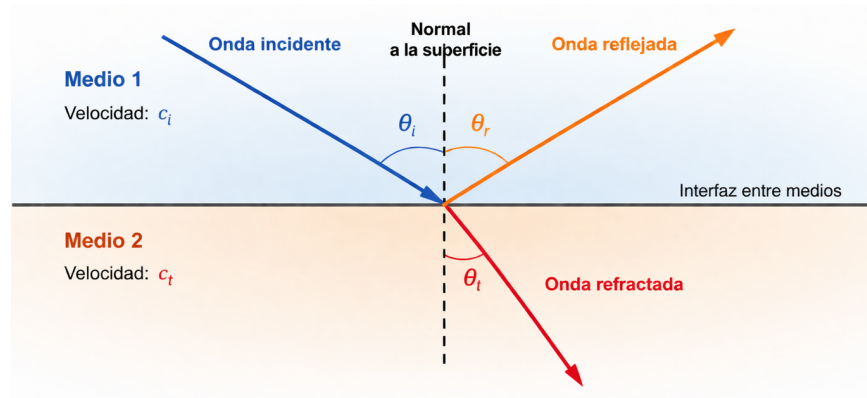


Figura 3.4: Interacción onda-interfaz

Difracción

Cuando una onda se enfrenta a un objeto o apertura de tamaño similar a su longitud de onda esta se ve expuesta a un fenomeno llamado difracción, el cual altera su dirección de propagación, causando una curvatura alrededor de estos obstaculos. A modo de ejemplo, es por la difracción que uno puede escuchar y mantener conversaciones con gente en una habitacion adyacente.

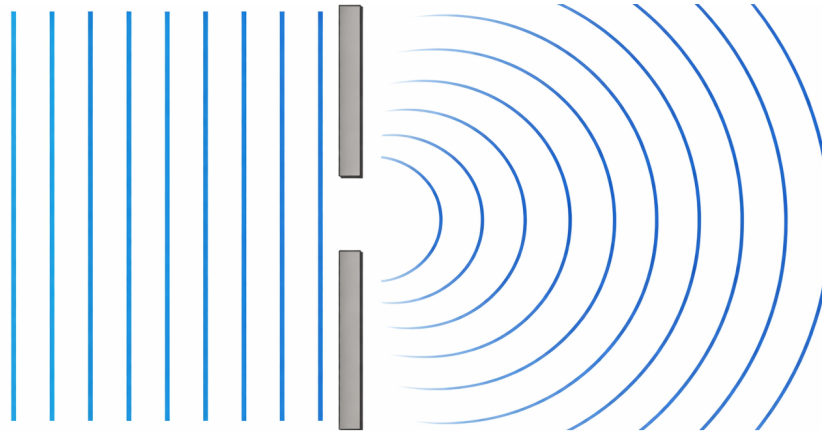


Figura 3.5: Difracción de onda

Atenuación

Uno de los fenomenos que se dan sin importar la existencia de una interfaz acústica es la atenuación. Se le llama atenuación a la disminución en la amplitud que experimenta una onda al propagarse a traves de un medio, producto de la capacidad de transformar la energia de la onda en calor y la dispersión que ocasionan discontinuidades de impedancia en un mismo medio que exagera la transformación de energia en calor.

3.1.4 Representaciones Tiempo-Frecuenciales

En primera instancia se suele enunciar situaciones que implican el movimiento oscilatorio de una partícula con representaciones en el dominio del tiempo, mediante funciones tipo $x(t) = f(t)$, donde se ilustra la posición espacial de la partícula para un tiempo dado. Esta representación temporal suele ser suficiente en estudios superficiales de vibraciones estacionarias que no presentan cambios en sus propiedades a traves del tiempo, sin embargo resulta inutil en el analisis de regimenes vibratorios mas complejos.

La primera aproximación a una herramienta para analizar vibraciones complejas es la serie armónica de Joseph Fourier, que permite representar una función periódica $f(t)$ de periodo T mediante una suma ponderada de senoidales de distintas frecuencias.

$$f(t) = \frac{1}{2}A_0 + A_1 \cos \omega t + A_2 \cos 2\omega t + \cdots + A_n \cos n\omega t + \cdots \\ + B_1 \sin \omega t + B_2 \sin 2\omega t + \cdots + B_n \sin n\omega t + \cdots \quad (3.4)$$

donde a través de la frecuencia fundamental $\omega = 2\pi/T$, y las constantes A y B se definen:

$$A_n = \frac{2}{T} \int_0^T f(t) \cos(n\omega t) dt \\ B_n = \frac{2}{T} \int_0^T f(t) \sin(n\omega t) dt \quad (3.5)$$

Transformada de Fourier

El principal problema de las series de Fourier es que solo aplica para funciones periódicas, sin embargo este problema se puede solucionar mediante un truco matemático haciendo $T \rightarrow \infty$, esto hace que la frecuencia fundamental $\omega \rightarrow 0$ permitiendo que los armónicos se evalúen de forma continua, representando así todas las frecuencias posibles.

Es así que a partir de este truco y la ecuación (3.4) nace la llamada transformada de Fourier y su inversa.

$$g(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \\ f(t) = \int_{-\infty}^{\infty} g(\omega) e^{j\omega t} d\omega \quad (3.6)$$

Donde $f(t)$ es una función no-periódica y $g(\omega)$ es su respectiva densidad espectral. En la primera ecuación se dice que $g(\omega)$ corresponde a la transformada de Fourier de $f(t)$, mientras que en la segunda ecuación se enuncia el caso inverso.

Transformada de Fourier Discreta

La transformada de Fourier clásica está diseñada para tratar con funciones definidas de manera continua en el dominio del tiempo, sin embargo la tecnología disponible en el mundo real no permite medir fenómenos ondulatorios en este formato. Las herramientas de medición actuales discretizan señales analógicas $x(t)$ definidas en el dominio temporal continuo, transformándolas en señales digitales $x[n]$ definidas en el dominio temporal discreto. Esta discretización es gobernada por el periodo de muestreo T_s , donde la señal analógica y digital se relacionan según $x[n] = x(nT_s)$.

Teniendo la función discretizada en el tiempo $x[n]$ a través de $N \in \mathbb{N}$ muestras, se procede a discretizar la frecuencia fundamental ω de la siguiente forma:

$$\omega_k = \frac{2\pi k}{NT_s} \quad (3.7)$$

Para finalmente llegar a la expresión de la DFT (Discrete Fourier Transform) con su respectiva inversa:

$$\begin{aligned} X[k] &= \sum_{n=0}^{N-1} x[n] e^{-i\frac{2\pi}{N}kn} \\ x[n] &= \frac{1}{N} \sum_{k=0}^{N-1} X[k] e^{i\frac{2\pi}{N}kn} \end{aligned} \quad (3.8)$$

Donde en la ecuación superior de 3.8 $X[k]$ corresponde a la transformada de Fourier discreta de $x[n]$, el caso inverso se enuncia en la ecuación inferior.

La DFT también puede ser expresada de forma matricial de la siguiente forma:

$$\begin{aligned} X &= Wx \\ W_{k,n} &= e^{-i\frac{2\pi}{N}kn} \end{aligned} \quad (3.9)$$

Esta representación de la transformada de Fourier discreta permite identificar fácilmente un costo operacional de $O(N^2)$. Esto cual conlleva una gran debilidad al momento de analizar señales de larga duración, por ejemplo, para una señal digital de 10^6 muestras se necesitarían un total de 10^{12} operaciones para calcular.

El proceso de cálculo de la DFT se puede acelerar mediante la FFT (Fast Fourier Transform), algoritmo que aprovecha de características como la simetría de complejo conjugado y periodicidad en n, k del factor $W_{k,n}$ definido en 3.9. Este algoritmo permite descomponer una DFT en otras 2 más pequeñas de forma sucesiva, permitiendo alcanzar un costo operacional de $O(N \log(N))$.

Transformada de Fourier de tiempo corto

La transformada de Fourier nos permite identificar que frecuencias senoidales componen una señal de interés, de manera global. Por ejemplo, podría indicar que un acorde La (compuesto por La4, Do5, Mi5) está compuesto principalmente por frecuencias de 440, 523 y 659 [Hz] principalmente. Al momento de representar una canción la transformada de Fourier entregaría las frecuencias de todos los acordes que la componen, sin embargo al aplicar transformada inversa resultaría un sonido correspondiente a las mismas frecuencias obtenidas siendo reproducidas al mismo tiempo. Esto deja en evidencia que al aplicar la transformada clásica se quiebra la estructura temporal de señales más complejas.

Para preservar la estructura temporal de las señales se propuso representarlas mediante Transformadas de Fourier evaluadas en periodos de tiempo sucesivos, bautizados "ventanas". Dando origen a la Transformada de Fourier de tiempo corto.

La Transformada de Fourier de tiempo corto discreta se enuncia en la siguiente expresión:

$$X[m, k] = \sum_{n=0}^{N-1} x[n + mH] w[n] e^{-i\frac{2\pi}{N}kn} \quad (3.10)$$

Donde m, k corresponden a los índices temporales y frecuenciales respectivamente, N al tamaño de la DFT individual (calculada mediante FFT), $w[n]$ a la función ventana que permite analizar intervalos de tiempo aislados dentro de la señal y H al salto temporal que se tiene entre ventanas de evaluación consecutivas.

3.2 Ultrasonido

El ultrasonido, en el ámbito físico, corresponde a todas las ondas mecánicas que superen el rango auditivo humano el cual supone ondas con frecuencias superiores a los 20 kHz. En un contexto médico, el ultrasonido corresponde a una de las técnicas de diagnóstico no-invasivas por excelencia, dado su amplia utilización producto de su altísima seguridad, fácil acceso y capacidad de entregar imágenes en tiempo real. La elaboración de imágenes por ultrasonido se basa en la emisión de un arreglo ondas ultrasónicas y posterior recepción de los ecos generados por la interacción de las ondas con las interfaces entre tejidos de distinta impedancia acústica. El análisis y post procesamiento de estos ecos permite confeccionar una imagen que representa la estructura interna de tejidos blandos.

3.2.1 Transductores

Tanto la emisión como recepción de ondas se basa en un fenómeno presente en algunos tipos de cristales llamado "efecto piezoeléctrico". Los cristales piezoeléctricos al ser sometidos a deformaciones generan una diferencia de potencial eléctrico, también se da el caso inverso donde al ser expuestos a una diferencia de potencial, estos se deforman. El instrumento que se encarga de emitir y recibir ondas se llama "transductor", dispositivo que cuenta con un arreglo de cristales piezoeléctricos los cuales pueden ser excitados por una diferencia de potencial para emitir ondas mecánicas de alta frecuencia y en el caso contrario pueden traducir las deformaciones que provocan los ecos, en señales eléctricas que permiten confeccionar imágenes para el diagnóstico.

Existen distintos tipos de transductores cuyas diferencias radican principalmente en la estructura que posea el arreglo de cristales piezoeléctricos que sostienen, en la figura 3.6 se muestran los tres tipos de transductores más comunes.

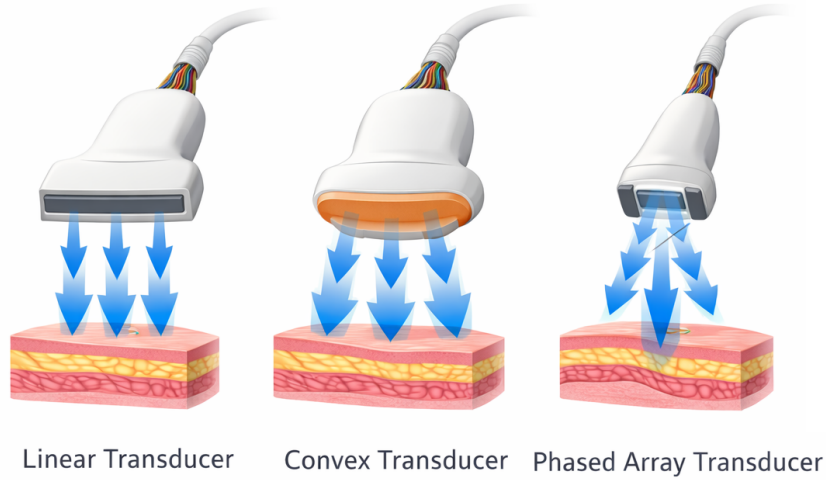


Figura 3.6: Transductor lineal, convexo y de matriz en fase

3.2.2 Transmisión y recepción

En esta etapa de la generación de imágenes de ultrasonido se diseñan las ondas que se propagarán a través del tejido, es por esto que el primer paso es determinar como será la excitación eléctrica a la que se verán expuestos los cristales piezoelectrónicos del transductor, por ejemplo, con un transductor lineal se podrían excitar todos los cristales al mismo tiempo para generar una onda ultrasónica plana.

La velocidad del sonido en tejidos corporales suele rondar $c_o = 1.5[mm/\mu s]$, teniendo en cuenta que el ultrasonido médico trabaja con ondas de $1 - 20[MHz]$ de frecuencia, utilizando la ecuación 3.1 podemos calcular que la longitud de onda que suelen tener las ondas al propagarse por el cuerpo es de $1.5[mm] - 75[\mu m]$. Considerando el tamaño de los objetos o "scatterers" a los que se enfrenta la onda se tiene una razón apertura-longitud de onda de a/λ en el orden de múltiplos de 10, esto quiere decir que las ondas ultrasónicas se propagan en un régimen difractivo a través del tejido [1].

Una vez que la onda ultrasónica emitida interactúa con el tejido de interés hasta reflejarse con un scatterer, este eco reflejado es captado por cada cristal piezoelectrónico del transductor y es transformado en una excitación eléctrica la cual es registrada en el equipo computacional en forma de series de tiempo de amplitud comúnmente llamada señal RF o radiofrecuencia.

3.2.3 Construcción imagen B-mode

Las señales RF captadas por cada uno de los canales del transductor representan ondas oscilatorias de alta frecuencia centradas cerca de la frecuencia central del transductor f_c , la cual usualmente ronda valores entre $1 - 40 [MHz]$. Cada una de estas señales pueden representarse mediante la siguiente expresión:

$$RF_i(t) = A_i(t) \cos(2\pi f_c t + \phi_i(t)) \quad (3.11)$$

Donde $RF_i(t)$ corresponde a la señal captada por el elemento de transductor numero i , $A_i(t)$ corresponde a la amplitud instantanea de la señal $RF_i(t)$, llamada academicamente como la envolvente. Dentro del coseno se hallan 2 factores, el primero $2\pi f_c t$ corresponde a la oscilación de alta frecuencia centrada en la banda basa del transductor, mientras que el segundo factor $\phi(t)$ representa la fase instantanea de la señal, es en este factor $\phi(t)$ donde esta almacenada información necesaria para construir una imagen B-mode coherente.

Para evitar trabajar con las altas oscilaciones de la frecuencia central del transductor, las señales RF pasan por un proceso de demodulación IQ, para ello se debe considerar la representación compleja de la señal RF presentada en primera instancia en la ecuación 3.11:

$$\begin{aligned} \theta(t) &= 2\pi f_c t + \phi(t) \\ RF_i(t) &= A_i(t) \cos(\theta(t)) + i A_i(t) \sin(\theta(t)) \\ RF_i(t) &= A_i(t) e^{i\theta(t)} \end{aligned} \quad (3.12)$$

La demodulación IQ es una tecnica de procesamiento de señales digitales en donde una señal portadora (en este caso la señal RF) puede ser representada por otras 2 señales ortogonales, es decir 2 señales que posean un desfase de $\pi/2$ entre ellas, bautizando la señal en fase como I y la seña en cuadratura como Q . Sabiendo que las funciones senoidales fundamentales estan en cuadratura por naturaleza, se aprovecha la nueva representación compleja de la señal RF de la siguiente manera:

$$\begin{aligned} RF_i(t) &= A_i(t) e^{i(2\pi f_c t + \phi(t))} \\ RF_i(t) \times e^{-i2\pi f_c t} &= A_i(t) e^{i(2\pi f_c t + \phi(t))} \times e^{-i2\pi f_c t} \\ IQ_i(t) &= A_i(t) e^{i\phi(t)} \end{aligned} \quad (3.13)$$

La demodulación IQ transformo la serie de tiempo de amplitud de la señal RF en una serie de tiempo de fasores (señal IQ). Ademas, como se puede ver en 3.13, la demodulación se deshizo de las oscilaciones de alta frecuencia de la banda base de transductor, dejando solo la información relevante que aportan $A(t)$ y $\phi(t)$.

Despues de haber demodulado las señales RF en señales IQ, se comienza el proceso de "beamforming delay and sum" o formación de haz retraso y suma. Este proceso consiste en sumar de forma coherente las señales recibidas por cada elemento del transductor para obtener la intensidad del eco de retorno recibido desde distintos puntos dentro del tejido de interés. Para ello, cada señal debe ser retrasada temporalmente según el tiempo de vuelo esperado entre el punto a evaluar y cada elemento del arreglo del transductor.

El primer paso de este procedimiento es establecer la grilla de pixeles que representarán la

región de interés sobre la cual se determinará la intensidad de los ecos de retorno. Cada pixel de la grilla corresponde a una posición $X_s = (x_s, z_s)$ dentro del tejido en la cual se asume la existencia de un reflector.

Una vez definida la grilla, se calcula, para cada pixel, el tiempo de vuelo de la onda ultrasónica en viajar desde el transductor hasta dicho pixel y volver nuevamente hacia cada elemento de transductor. Este tiempo de vuelo depende de la geometría del sistema y de la velocidad de propagación del sonido en el medio estudiado. Según Perrot et al. [2], el tiempo de vuelo bidireccional puede expresarse como:

$$\tau_i(X_s) = \frac{d_{TX}(X_s) + d_{RX}(X_s, x_i)}{c} - t_0 \quad (3.14)$$

Donde d_{TX} corresponde a la distancia recorrida durante la etapa de transmisión, d_{RX} la distancia recorrida entre el reflector y el elemento i receptor del transductor, c es la velocidad del sonido en el medio estudiado y t_0 representa el tiempo inicial de adquisición.

En emisión de ondas planas no-inclinadas con transductores lineales, la distancia de transmisión corresponde a $d_{TX} = z_s$, mientras que la distancia de recepción d_{RX} puede calcularse mediante:

$$d_{RX}(X_s, x_i) = \sqrt{(x_i - x_s)^2 + z_s^2} \quad (3.15)$$

Una vez calculados los tiempos de vuelo para cada pixel y cada elemento receptor, se procede a evaluar la señal IQ_i en el instante $\tau_i(X_s)$ correspondiente. Dado que los tiempos de vuelo son valores continuos y las señales están discretizadas, los valores $IQ_i(\tau_i(X_s))$ generalmente no coinciden con ninguna muestra almacenada, por lo que deben estimarse mediante interpolación a partir de las muestras adyacentes. Sin embargo, dado que la suma y la demodulación IQ son operadores no conmutativos, sumar directamente las señales IQ_i retrasadas no preserva la coherencia de fase entre ellas. Para corregir esto, es necesario aplicar un rotador de fase a cada señal antes de sumarla [2]. La señal beamformada en el pixel X_s se obtiene entonces como:

$$IQ_{bf}(X_s) = \sum_{i=1}^{N_e} IQ_i(\tau_i(X_s)) e^{j2\pi f_c \tau_i(X_s)} \quad (3.16)$$

donde el factor $e^{j2\pi f_c \tau_i(X_s)}$ es el rotador de fase que compensa el desfase introducido al retrasar la señal IQ en el dominio del tiempo, y N_e es el número de elementos del transductor que participan en la suma. La señal $IQ_{bf}(X_s)$ es compleja, con parte real e imaginaria correspondiente a las componentes en fase y en cuadratura de la señal beamformada.

Completado el proceso de beamforming para todos los pixeles de la grilla, se obtiene una matriz de valores complejos $IQ_{bf}(X_s)$. Para construir la imagen B-mode, se calcula la envolvente real de la señal beamformada, definida como el módulo de la señal compleja:

$$E(X_s) = |IQ_{bf}(X_s)| = \sqrt{I_{bf}^2(X_s) + Q_{bf}^2(X_s)} \quad (3.17)$$

donde I_{bf} y Q_{bf} son las partes real e imaginaria de IQ_{bf} , respectivamente. La envolvente $E(X_s)$ representa la amplitud instantánea del eco de retorno en cada punto del tejido.

Finalmente, dado que la amplitud de los ecos en tejido biológico puede variar en varios órdenes de magnitud entre distintas estructuras, se aplica una compresión logarítmica para adaptar el rango dinámico de la señal a la capacidad de visualización del sistema y a la percepción visual humana. La imagen B-mode se obtiene como:

$$B(X_s) = 20 \log_{10} \left(\frac{E(X_s)}{E_{\max}} \right) \quad (3.18)$$

donde $E_{\max}$ corresponde al valor máximo de la envolvente en la imagen, de modo que los valores de $B(X_s)$ quedan expresados en decibels (dB) con un máximo de 0 dB. Las regiones del tejido con alta reflectividad aparecen cercanas a 0 dB (brillantes), mientras que las regiones con baja reflectividad, como cavidades quísticas, toman valores muy negativos (oscuras). El rango dinámico típicamente visualizado en imágenes B-mode clínicas es de entre 40 y 60 dB.

3.3 Transporte óptimo

El transporte óptimo es una rama de análisis matemático enfocada en la transformación óptima entre distribuciones de probabilidad minimizando el costo asociado al transporte. Esta teoría remonta a Francia del siglo XVIII, cuando la École Royale le encomienda al matemático Gaspard Monge idear un plan de construcción de fortificaciones militares que permitiera edificarlas rápidamente sin comprometer una defensa robusta.

Si bien la formulación de Monge permitio formalizar los cimientos de la teoría de transporte óptimo esta tenia una estructura difícil de analizar. La investigación de este problema se estanco por 150 años hasta que el economista soviético Leonid Kantorovich, en el año 1942, planteo una relajación de la formulación de Monge, transformando el problema en uno convexo mucho más fácil de tratar, revitalizando el desarrollo de esta disciplina.

En la época moderna destaca el matemático Yann Bernier quien demostro la existencia de una respuesta a la formulación de Monge con un costo cuadrático asociado la cual corresponderia al gradiente de una función convexa. Para efectos de esta memoria el personaje más importante es Marco Cuturi quien propuso resolver una formulación de Kantorovich con regularización entrópica, llegando a un problema estrictamente convexo fácil de resolver de manera computacional.

3.3.1 Formulación de Monge

El transporte óptimo tiene sus raíces en la publicación de Gaspard Monge titulada "Mémoire sur la théorie des déblais et des remblais" [3] en el año 1781, donde aborda un problema de optimización en asignación de recursos de una forma geométrica. Si bien Monge planteó correctamente el problema desde un punto de vista físico, carecía de las herramientas matemáticas necesarias para formalizarlo, especialmente una forma de describir de una forma general los conjuntos que él buscaba transportar de forma óptima. La teoría de la medida otorga una herramienta muy fuerte que permite expresar de forma matemática el "tamaño" o "cantidad" de un conjunto, permitiendo así generalizar conceptos como el área, volumen o masa mediante el concepto de "medida".

El transporte óptimo entrega una herramienta muy útil para comparar medidas μ y ν desde una perspectiva energética. Asumiendo que ambas medidas son de probabilidad las cuales viven en los espacios X e Y respectivamente, se define una función costo $c : X \times Y \rightarrow [0, +\infty]$ donde $c(x, y)$ cuantiza el costo de transportar una unidad de masa en la posición $x \in \mu$ a otra $y \in \nu$. El problema de transporte óptimo consiste en encontrar la forma de transportar μ hacia ν que minimice el valor de c . Para ello se debe dar una definición formal del concepto de "transporte".

Definición 3.3.1. Se dice que $T : X \rightarrow Y$ transporta $\mu \in \mathcal{P}(X)$ hacia $\nu \in \mathcal{P}(Y)$, y llamamos T "un mapa de transporte", si:

$$\nu(B) = \mu(T^{-1}(B)) \quad \text{para todo set } B \text{ } \nu\text{-medible.} \quad (3.19)$$

Esta definición sostiene que para cualquier región $B \subseteq Y$, la medida de probabilidad ν debe asignar a la misma "masa" que la medida de probabilidad μ asigna a la región $T^{-1}(B) \subseteq X$ correspondiente a la región original de X que fue trasladada según T hacia B , asegurando conservación de la masa. Usualmente la ecuación 3.19 se abrevia $\nu = T_{\#}\mu$ donde ν es el "pushforward" de μ según T , notación que se utiliza para definir el problema de optimización de Monge.

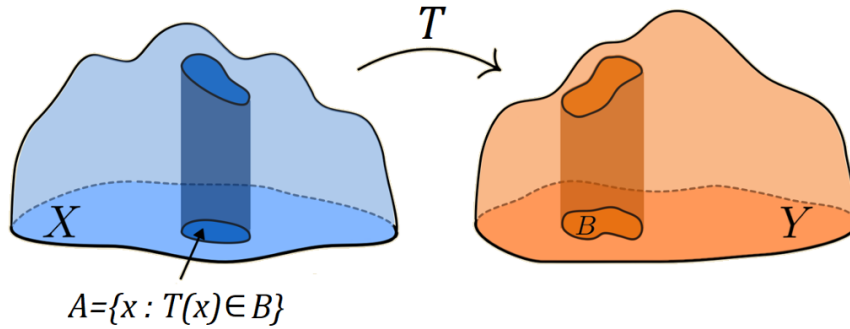


Figura 3.7: Ilustración del mapa de transporte de Monge, ν es el pushforward de μ según T

Definición 3.3.2. Formulación de Monge del Problema de Transporte Óptimo:

Dado $\mu \in \mathcal{P}(X)$ y $\nu \in \mathcal{P}(Y)$,

$$\min_T \left\{ \int_X c(x, T(x)) d\mu(x) \mid T_{\#}\mu = \nu \right\} \quad (3.20)$$

Esta definición permite expresar de forma matemática lo que se estableció al principio de esta subsección. Este problema de optimización busca el plan de transporte $T : X \rightarrow Y$, que cumpla con la condición que ν sea el "pushforward" de μ , el cual minimice la integral. El término $c(x, T(x))$ corresponde al costo de transportar una unidad de masa desde una posición $x \in X$ hasta otra $T(x) \in Y$, por lo tanto el valor de la integral es el costo total de trasladar toda la distribución de masa que define μ sobre el espacio X hacia la distribución que ν define sobre el espacio Y según el mapa de transporte T .

Si se asume que T es biyectivo, T^{-1} es diferenciable y las distribuciones de probabilidad μ y ν son absolutamente continuas con respecto a la medida de Lebesgue en $\mathbb{R}^n$, siendo capaz de describirlas con funciones de densidad $d\mu(x) = f(x) dx$ y $d\nu(y) = g(y) dy$, se puede demostrar que la restricción (3.19) de conservación de la masa es equivalente a:

$$f(x) = g(T(x)) \left| \det(\nabla T(x)) \right|. \quad (3.21)$$

Esta nueva restricción es altamente no lineal, dificultando su tratamiento con herramientas de cálculo que permitan demostrar existencia y regularidad del mapa de transporte T , sin embargo esta no es su principal debilidad.

La formulación de Monge 3.20 tiene una limitación importante a raíz de como está definido el transporte T entre distribuciones. El mapa T asigna a cada punto x un único destino $T(x)$, por lo que no permite dividir la masa asociada a un punto y distribuirla en múltiples ubicaciones, provocando que el transporte se restrinja a correspondencias uno a uno.

3.3.2 Formulación de Kantorovich

El principal aporte de Kantorovich [4] fue relajar el carácter determinístico de los mapas de transporte de Monge, es decir, el hecho de que la masa ubicada en la posición x solo puede ser trasladada a una única posición $T(x)$. Kantorovich propone reemplazar los mapas de transporte T por una medida conjunta $\pi \in \mathcal{P}(X \times Y)$, llamada plan de transporte.

Definición 3.3.3. Dado $\mu \in \mathcal{P}(X)$ y $\nu \in \mathcal{P}(Y)$, se dice que $\pi \in \mathcal{P}(X \times Y)$ es un plan de transporte si:

$$\begin{aligned} \pi(A \times Y) &= \mu(A), \quad \text{para todo set } A \text{ } \mu\text{-medible,} \\ \pi(X \times B) &= \nu(B), \quad \text{para todo set } B \text{ } \nu\text{-medible.} \end{aligned} \quad (3.22)$$

Se le llama conjunto de planes de transporte admisibles a $\Pi(\mu, \nu)$. Este es compuesto por todos los planes de transporte que cumplan con 3.22, es decir, que su primer y segundo marginal correspondan con las medidas μ y ν respectivamente.

Esta nueva forma de expresar el transporte mediante la medida conjunta $\pi(x, y)$ permite que la masa asociada a un punto x de la distribución inicial μ sea transportada a múltiples puntos destino en la distribución final ν . En efecto, para un mismo valor de x existen múltiples valores de y para los cuales $\pi(x, y_i) > 0$, lo cual implica que la masa originalmente concentrada en x se divide en múltiples destinos y_i .

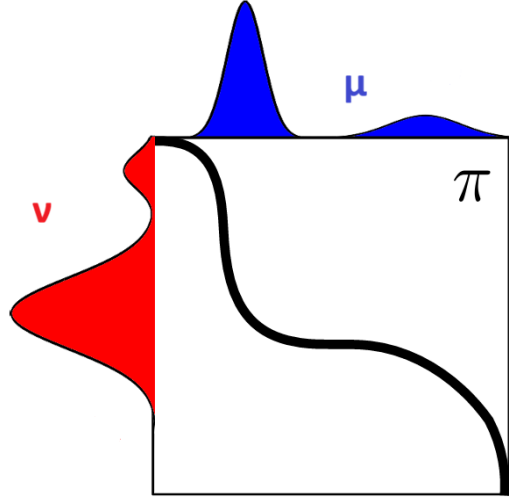


Figura 3.8: Plan de transporte π entre distribuciones continuas μ y ν

Definición 3.3.4. Formulación de Kantorovich del Problema de Transporte Óptimo:
Dado $\mu \in \mathcal{P}(X)$ y $\nu \in \mathcal{P}(Y)$,

$$\min_{\pi \in \Pi(\mu, \nu)} \left\{ \int_{X \times Y} c(x, y) d\pi(x, y) \right\} \quad (3.23)$$

La función costo $c(x, y)$ cuantiza el trabajo necesario para transportar una unidad de masa desde una posición x hasta otra posición y . Notar que de forma similar a la medida conjunta $\pi(x, y)$, $c(x, y)$ permite cuantizar el transporte desde una única posición x a distintos destinos y_i . Esto quiere decir que el problema de optimización descrito en 3.23 busca entre el conjunto de planes de transporte admisibles $\Pi(\mu, \nu)$ un plan de transporte admisible $\pi(x, y)$ que permita minimizar el valor de la integral, la cual representa el costo de transportar toda la masa que describe la medida conjunta $\pi(x, y)$ sobre la región $X \times Y$.

Si bien la formulación anterior está planteada sobre las medidas continuas μ y ν , en la aplicación práctica de la teoría de transporte óptimo resulta conveniente considerar una versión discreta del

problema, algo que resultaba imposible con la formulación de Monge. Es por esto que, a partir de ahora, se asume que las distribuciones de masa se encuentran concentradas sobre un número finito de puntos:

$$\mu = \sum_{i=1}^n a_i \delta_{x_i}, \quad \nu = \sum_{j=1}^m b_j \delta_{y_j}. \quad (3.24)$$

Donde $x_i \in X$ y $y_j \in Y$ representan posiciones discretas, a_i y b_j corresponden a las masas asociadas, δ_x y δ_y denotan medidas de Dirac centradas en x e y respectivamente.

Bajo esta nueva formulación discreta, el problema de transporte óptimo puede reformularse en términos de una representación matricial de los planes de transporte.

$$\Pi = (\pi_{ij}) \in \mathbb{R}^{n \times m}, \quad (3.25)$$

Donde cada elemento π_{ij} representa la cantidad de masa transportada desde la posición x_i hacia y_j .

De esta manera, el problema planteado de forma continua en 3.23 puede ser expresado de forma discreta de la siguiente manera:

Definición 3.3.5. *Formulación discreta de Kantorovich del Problema de Transporte Óptimo:*

Dado $\mu \in \mathcal{P}(X)$ y $\nu \in \mathcal{P}(Y)$,

$$\min_{\pi \in \Pi(\mu, \nu)} \left\{ \sum_{i,j} c_{ij} \pi_{ij} \mid \sum_j \pi_{ij} = a_i, \sum_i \pi_{ij} = b_j \right\} \quad (3.26)$$

Esta nueva forma de expresar el problema de transporte óptimo facilita la explicación de otra propiedad importante que posee el problema de Kantorovich, la cual es su capacidad de admitir una formulación dual. Esta dualidad resulta especialmente importante tanto desde el punto de vista teórico como computacional, ya que permite reinterpretar el problema en términos de potenciales asociados a las distribuciones de origen y destino.

Para obtener la formulación dual del problema de Kantorovich, se introduce el lagrangiano asociado a las restricciones marginales del problema primal 3.26. Para ello, se consideran los multiplicadores de Lagrange f_i y g_j , asociados respectivamente a las restricciones:

$$\sum_j \pi_{ij} = a_i, \quad \sum_i \pi_{ij} = b_j. \quad (3.27)$$

De esta forma el Lagrangiano puede escribirse:

$$\mathcal{L}(\Pi, f, g) = \sum_{i,j} c_{ij} \pi_{ij} + \sum_i f_i \left(a_i - \sum_j \pi_{ij} \right) + \sum_j g_j \left(b_j - \sum_i \pi_{ij} \right). \quad (3.28)$$

$$\mathcal{L}(\Pi, f, g) = \sum_i f_i a_i + \sum_j g_j b_j + \sum_{i,j} \pi_{ij} (c_{ij} - f_i - g_j). \quad (3.29)$$

Sabiendo que los multiplicadores de Lagrange f_i y g_j actúan como penalizadores que castigan planes de transporte π que no cumplan con las restricciones establecidas en 3.27, el problema primal 3.26 puede expresarse como:

$$\min_{\Pi \geq 0} \max_{f, g} \mathcal{L}(\Pi, f, g). \quad (3.30)$$

Dado que $\pi_{ij} \geq 0$, si el término $(c_{ij} - f_i - g_j) < 0$ puede hacer que el lagrangiano diverja a $-\infty$. Entonces para que el ínfimo sea finito, debe cumplirse:

$$f_i + g_j \leq c_{ij}, \quad \forall i, j. \quad (3.31)$$

Bajo esta condición, el mínimo respecto a Π se alcanza tomando $\pi_{ij} = 0$ en el último término, obteniendo:

$$\inf_{\Pi \geq 0} \mathcal{L}(\Pi, f, g) = \sum_i f_i a_i + \sum_j g_j b_j. \quad (3.32)$$

Definición 3.3.6. *Formulación discreta del problema dual de Kantorovich:*

Dado $\mu = \sum_{i=1}^n a_i \delta_{x_i}$ y $\nu = \sum_{j=1}^m b_j \delta_{y_j}$,

$$\max_{f, g} \left\{ \left(\sum_i f_i a_i + \sum_j g_j b_j \right) \mid f_i + g_j \leq c_{ij} \right\} \quad (3.33)$$

A diferencia del problema original, donde las variables desconocidas correspondían al plan de transporte π , en la formulación dual corresponden a f_i y g_j las cuales representan potenciales asociados a las distribuciones μ y ν respectivamente. La restricción presentada por primera vez en 3.31 impone que la suma de ambos potenciales no puede superar el costo de transportar una unidad de masa desde x_i a y_j , provocando que los potenciales generen una cota inferior del costo total de transporte. Por lo tanto, la formulación dual de transporte óptimo busca la mayor cota inferior del costo total de transporte. Bajo la suposición de costos finitos definidos por c e igualdad en las masas de las distribuciones iniciales y finales, el valor al cual llega la formulación dual coincide con el de la formulación primal.

3.3.3 Distancias y barycentros de Wasserstein

La formulación de Kantorovich permite definir de forma natural una familia de métricas sobre el espacio de medidas de probabilidad, conocidas como distancias de Wasserstein. Estas métricas cuantifican el costo mínimo necesario para transportar una distribución μ hacia otra ν , y constituyen una de las herramientas mas interesantes que nacen a partir del análisis de problemas de transporte óptimo.

Definición 3.3.7 (Distancia de Wasserstein de orden p). *Sea $p \geq 1$ y sean $\mu, \nu \in \mathcal{P}(X)$ medidas de probabilidad sobre un espacio métrico (X, d) . La distancia de Wasserstein de orden p entre μ y ν se define como*

$$W_p(\mu, \nu) = \left(\inf_{\pi \in \Pi(\mu, \nu)} \int_{X \times X} d(x, y)^p d\pi(x, y) \right)^{1/p}, \quad (3.34)$$

donde $\Pi(\mu, \nu)$ denota el conjunto de planes de transporte admisibles definido en 3.25.

El valor dentro del ínfimo corresponde exactamente al costo del problema planteado por Kantorovich en 3.23 con función costo $c(x, y) = d(x, y)^p$. De esta manera, $W_p(\mu, \nu)^p$ representa el costo de transporte óptimo entre las distribuciones μ y ν bajo el costo de potencia p , y $W_p(\mu, \nu)$ normaliza dicho valor para recuperar las propiedades de una métrica. Es decir, W_p satisface:

1. **No negatividad:** $W_p(\mu, \nu) \geq 0$, con $W_p(\mu, \nu) = 0$ si y solo si $\mu = \nu$.
2. **Simetría:** $W_p(\mu, \nu) = W_p(\nu, \mu)$.
3. **Desigualdad triangular:** $W_p(\mu, \gamma) \leq W_p(\mu, \nu) + W_p(\nu, \gamma)$ para toda $\gamma \in \mathcal{P}_p(X)$.

Proof. La no negatividad es inmediata de la definición. Para la identidad de los indiscernibles, nótese que el plan $\pi = (\text{id}, \text{id})_{\#}\mu$ (la medida concentrada en la diagonal) pertenece a $\Pi(\mu, \mu)$ y tiene costo cero, luego $W_p(\mu, \mu) = 0$. Recíprocamente, si $W_p(\mu, \nu) = 0$ entonces existe una sucesión de planes $\pi_k \in \Pi(\mu, \nu)$ con $\int d(x, y)^p d\pi_k \rightarrow 0$, lo que implica que π_k converge débilmente a un plan concentrado en la diagonal, forzando $\mu = \nu$.

La simetría se obtiene observando que si $\pi \in \Pi(\mu, \nu)$ entonces el plan imagen bajo la transposición $(x, y) \mapsto (y, x)$ pertenece a $\Pi(\nu, \mu)$ y tiene el mismo costo, ya que $d(x, y) = d(y, x)$.

La desigualdad triangular se demuestra mediante el *gluing lemma*: dados $\pi_1 \in \Pi(\mu, \nu)$ y $\pi_2 \in \Pi(\nu, \gamma)$ óptimos, existe una medida de probabilidad $\tilde{\pi}$ sobre X^3 con marginales (π_1, π_2) sobre los pares (x, y) e (y, z) respectivamente. Aplicando la desigualdad triangular de d y la desigualdad de Minkowski para integrales se obtiene

$$W_p(\mu, \gamma) \leq \left(\int d(x, z)^p d\tilde{\pi} \right)^{1/p} \leq \left(\int d(x, y)^p d\pi_1 \right)^{1/p} + \left(\int d(y, z)^p d\pi_2 \right)^{1/p} = W_p(\mu, \nu) + W_p(\nu, \gamma).$$

□

El hecho de que W_p sea una métrica produce que el espacio $\mathcal{P}_p(X)$ de medidas de probabilidad con momento de orden p finito adquiera la estructura de un espacio métrico, en donde cada medida de probabilidad puede ser interpretada como un punto o ubicación en ese espacio, mientras que $W_p(\mu, \nu)$ representa la distancia entre ambas distribuciones.

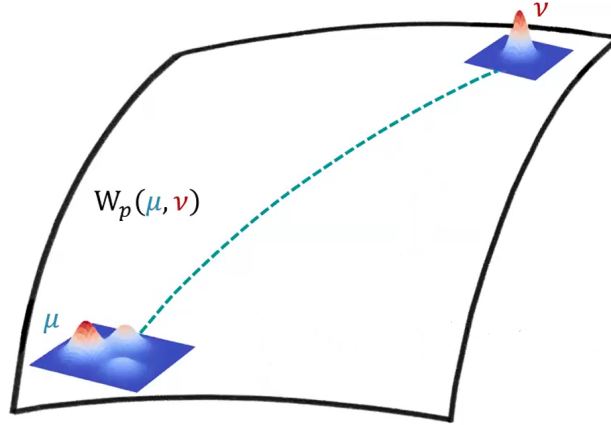


Figura 3.9: Ilustración del espacio métrico que W_p define sobre el espacio de medidas de probabilidad de momento de orden p finito

Esta métrica W_p definida sobre el espacio de medidas de probabilidad con momento de orden p finito, permite inducir una noción de distancia intermedia o promedio entre dos distribuciones de probabilidad. Esta noción se conoce como *baricentro de Wasserstein*, la cual generaliza el concepto de media ponderada al espacio de medidas de probabilidad.

Definición 3.3.8 (Baricentro de Wasserstein). Sean $\mu_1, \dots, \mu_K \in \mathcal{P}_p(X)$ medidas de probabilidad y sean $\lambda_1, \dots, \lambda_K \geq 0$ pesos tales que $\sum_{k=1}^K \lambda_k = 1$. El baricentro de Wasserstein de orden p con respecto a los pesos $\{\lambda_k\}$ se define como la medida $\mu^* \in \mathcal{P}_p(X)$ que resuelve el problema de optimización

$$\mu^* = \min_{\mu \in \mathcal{P}_p(X)} \sum_{k=1}^K \lambda_k W_p^p(\mu, \mu_k). \quad (3.35)$$

El baricentro de Wasserstein busca la distribución μ^* que minimiza la suma ponderada de los costos de transporte hacia cada una de las distribuciones $\mu_1, \dots, \mu_K$. De esta forma, μ^* representa la distribución de probabilidad que se encuentra en la posición más "central" respecto a las otras distribuciones dadas, bajo la geometría inducida por W_p .

Si bien el baricentro de Wasserstein establece un concepto de promedio en el espacio de medidas de probabilidad, existe una forma alternativa y más general de construir distribuciones interpoladas entre dos medidas de probabilidad μ y ν , la cual puede aplicarse si se conoce el plan de transporte π entre ambas.

Definición 3.3.9 (Interpolación mediante plan de transporte). Sean $\mu_1, \mu_2 \in \mathcal{P}_p(X)$ y sea $\pi \in \Pi(\mu_1, \mu_2)$ un plan de transporte con marginales μ_1 y μ_2 . Para $\alpha \in [0, 1]$, se define el mapa de interpolación $T_\alpha : X \times X \rightarrow X$ como

$$T_\alpha(x, y) = (1 - \alpha)x + \alpha y. \quad (3.36)$$

La distribución interpolada con parámetro α inducida por π se define como la imagen de π bajo T_α ,

$$\mu_\alpha = (T_\alpha)_\# \pi, \quad (3.37)$$

es decir, μ_α es la distribución de la variable aleatoria $(1 - \alpha)X + \alpha Y$, donde $(X, Y) \sim \pi$.

A diferencia de la interpolación lineal $(1 - \alpha)\mu_1 + \alpha\mu_2$, que mezcla ambas densidades sin tener en cuenta la geometría de las distribuciones entre las que se interpola, la interpolación inducida por el plan de transporte $\pi \in \Pi(\mu_1, \mu_2)$ desplaza masas de forma coherente. En particular, la curva μ_α constituye una geodésica en el espacio de distribuciones de probabilidad, dando lugar a una noción de "camino más corto" entre distribuciones.

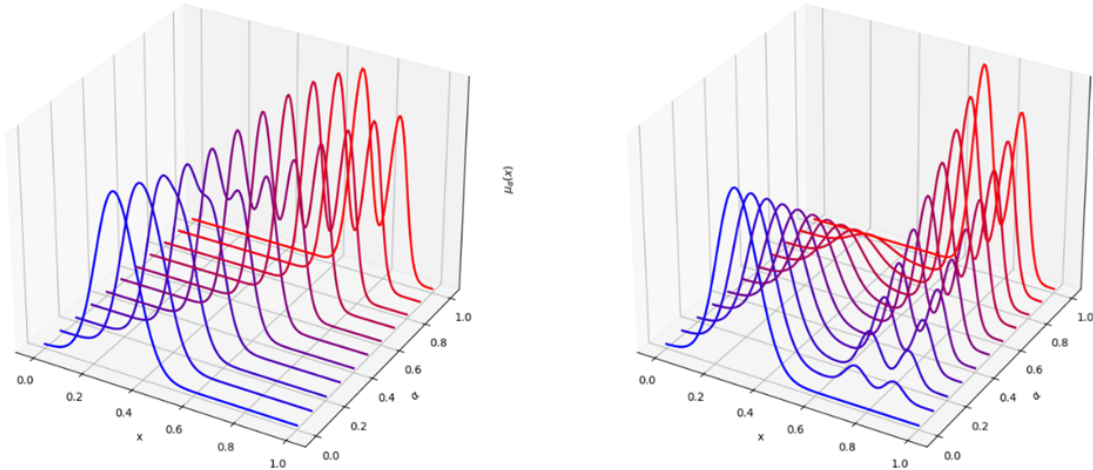


Figura 3.10: Comparación de interpolación inducida por plan de transporte óptimo π con interpolación lineal para un mismo par de distribuciones de probabilidad

3.3.4 Transporte óptimo unidimensional

El caso unidimensional del transporte óptimo admite una solución explícita y elegante la cual no se hereda a dimensiones superiores. Esta nace a partir de una observación clave sobre la estructura del plan de transporte óptimo: la masa no puede realizar cruces. Esta propiedad conocida como monotonicidad permite identificar el mapa de transporte óptimo de manera cerrada en términos de las funciones de distribución acumulada (CDF) de μ y ν .

A lo largo de esta subsección se asume que $X = Y = \mathbb{R}$, que μ y ν son medidas de probabilidad sobre $\mathbb{R}$, y que la función de costo es $c(x, y) = |x - y|^p$ para algún $p \geq 1$.

Definición 3.3.10 (Propiedad de monotonicidad en una dimensión). *Sea $\pi \in \Pi(\mu, \nu)$ un plan de transporte óptimo para el costo $c(x, y) = |x - y|^p$ con $p \geq 1$. Entonces π no presenta cruces, es decir, para $\pi \otimes \pi$ -casi todo par de puntos $(x_1, y_1), (x_2, y_2) \in \text{sop}(\pi)$ se cumple:*

$$(x_1 - x_2)(y_1 - y_2) \geq 0.$$

En particular, $x_1 \leq x_2$ implica $y_1 \leq y_2$.

La propiedad de monotonicidad puede demostrarse por contradicción.

Proof. Supóngase que existen $(x_1, y_1), (x_2, y_2) \in \text{sop}(\pi)$ con $x_1 < x_2$ pero $y_1 > y_2$. Construyase un plan alternativo $\tilde{\pi}$ reemplazando localmente la masa transportada entre estos cuatro puntos: en lugar de enviar masa desde x_1 hacia y_1 y desde x_2 hacia y_2 , se envía masa desde x_1 hacia y_2 y desde x_2 hacia y_1 . Formalmente, para $\varepsilon > 0$ suficientemente pequeño, se define

$$\tilde{\pi} = \pi - \varepsilon (\delta_{(x_1, y_1)} + \delta_{(x_2, y_2)}) + \varepsilon (\delta_{(x_1, y_2)} + \delta_{(x_2, y_1)}).$$

Este plan $\tilde{\pi}$ pertenece a $\Pi(\mu, \nu)$ pues sus marginales coinciden con las de π . La diferencia de costos entre π y $\tilde{\pi}$ es

$$\Delta = \varepsilon [|x_1 - y_1|^p + |x_2 - y_2|^p - |x_1 - y_2|^p - |x_2 - y_1|^p].$$

Para $p \geq 1$, la convexidad de $c(x, y) = |x - y|^p$ implica que para cualesquiera reales a, b, c, d con $a < b$ y $c > d$,

$$|a - c|^p + |b - d|^p > |a - d|^p + |b - c|^p,$$

lo cual se verifica aplicando la desigualdad de convexidad estricta al segmento entre los pares $(a - c, b - d)$ y $(a - d, b - c)$. En consecuencia $\Delta > 0$, de modo que $\tilde{\pi}$ tiene un costo estrictamente menor que π , contradiciendo la optimalidad de π . Por lo tanto, no pueden existir tales pares con cruce, y la propiedad queda demostrada. $\square$

La propiedad de monotonicidad caracteriza completamente al mapa de transporte entre μ y ν , el cual es definido a través de las inversas generalizadas de sus CDFs. Recordar que la función de distribución acumulada de una medida μ sobre $\mathbb{R}$ se define como $F_\mu(x) = \mu((-\infty, x])$, y su inversa generalizada (o función cuantil) como:

$$F_\mu^{-1}(t) = \inf\{x \in \mathbb{R} : F_\mu(x) \geq t\}, \quad t \in (0, 1).$$

Definición 3.3.11 (Solución explícita del transporte óptimo en 1D). Sean $\mu, \nu \in \mathcal{P}(\mathbb{R})$ y sea $c(x, y) = |x - y|^p$ con $p \geq 1$. Si μ es absolutamente continua con respecto a la medida de Lebesgue, entonces el problema de Kantorovich 3.26 admite un único plan óptimo, el cual es inducido por el mapa de transporte

$$T^* = F_\nu^{-1} \circ F_\mu, \quad (3.35)$$

es decir, $\pi^* = (\text{id}, T^*)_{\#} \mu$. Este mapa es monótono no decreciente y satisface $T^*_{\#} \mu = \nu$.

Proof. Paso 1: T^* es un mapa de transporte válido. Sea $U \sim \text{Unif}(0, 1)$. Por la propiedad de la transformada integral de probabilidad, $F_\mu^{-1}(U) \sim \mu$ y $F_\nu^{-1}(U) \sim \nu$. Por lo tanto, el par $(F_\mu^{-1}(U), F_\nu^{-1}(U))$ define un acoplamiento entre μ y ν , cuya ley es el plan $\pi^* = (\text{id}, T^*)_{\#} \mu \in \Pi(\mu, \nu)$. La monotonicidad de T^* es consecuencia directa de que F_μ^{-1} y F_ν^{-1} son ambas funciones no decrecientes.

Paso 2: π^* es óptimo. Sea $\pi \in \Pi(\mu, \nu)$ cualquier plan de transporte. Para cualesquiera funciones medibles f y g , la desigualdad de Young implica que para $p > 1$ con exponente conjugado $q = p/(p - 1)$,

$$|x - y|^p \geq f(x) + g(y) \quad \text{siempre que} \quad f(x) + g(y) \leq |x - y|^p.$$

En el caso 1D con costo $|x - y|^p$, se puede construir un par de funciones (f^*, g^*) que satisfacen la condición de dualidad de Kantorovich $f^*(x) + g^*(y) \leq |x - y|^p$ para todo (x, y) , con igualdad sobre $\text{sop}(\pi^*)$. Integrando,

$$\int |x - y|^p d\pi(x, y) \geq \int f^* d\mu + \int g^* d\nu = \int |x - T^*(x)|^p d\mu(x),$$

donde la última igualdad se obtiene porque la igualdad dual se alcanza en π^* . Esto muestra que π^* minimiza el costo sobre todos los planes en $\Pi(\mu, \nu)$.

Paso 3: Unicidad. Como μ es absolutamente continua, el mapa $T^* = F_\nu^{-1} \circ F_\mu$ está definido μ -casi en todo punto. La propiedad de monotonicidad garantiza que todo plan óptimo debe ser no cruzado, y puesto que μ no tiene átomos, el único plan no cruzado en $\Pi(\mu, \nu)$ es aquel inducido por un mapa monótono. La unicidad μ -c.s. del mapa cuantil $F_\nu^{-1} \circ F_\mu$ concluye el argumento. $\square$

3.3.5 Formulación de Kantorovich regularizada entrópicamente

Si bien la formulación original de Kantorovich expuesta en 3.26 constituye una relajación del problema de Monge 3.20 que garantiza la existencia de una solución bajo hipótesis adecuadas, su resolución numérica continúa siendo costosa para problemas de gran escala. En particular, la búsqueda del plan de transporte óptimo requiere resolver un problema de programación lineal sobre

una matriz de transporte cuyas dimensiones crecen cuadráticamente con el número de puntos que discretizan las distribuciones involucradas.

Con el objetivo de obtener algoritmos más eficientes, Marco Cuturi en el año 2013 [5] propuso incorporar un término de regularización entrópica a la formulación de Kantorovich. Esta modificación favorece planes de transporte más suaves y distribuidos, llegando a una versión relajada del problema original cuya estructura matemática permite el desarrollo de algoritmos iterativos altamente eficientes. A continuación se presenta la formulación discreta del problema de Kantorovich regularizado entrópicamente:

Definición 3.3.12. *Formulación discreta de Kantorovich con regularización entrópica:*

Dado $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$ y ϵ

$$\min_{\pi \in \Pi(\mu, \nu)} \left\{ \sum_{i,j} c_{ij} \pi_{ij} - \epsilon H(\pi) \mid \sum_j \pi_{ij} = a_i, \sum_i \pi_{ij} = b_j \right\} \quad (3.34)$$

Donde $H(\pi)$ corresponde a la entropía de Shannon asociada al plan de transporte π , definida de la siguiente manera:

$$H(\pi) = - \sum_{i,j} \pi_{ij} (\log \pi_{ij} - 1) \quad (3.35)$$

Esta nueva formulación del problema de transporte óptimo permite construir un algoritmo iterativo el cual nace a partir de la optimización mediante el Lagrangiano.

$$\min_{\pi \in \Pi(\mu, \nu)} \left\{ \sum_{i,j} c_{ij} \pi_{ij} + \epsilon \sum_{i,j} \pi_{ij} (\log \pi_{ij} - 1) \mid \sum_j \pi_{ij} = a_i, \sum_i \pi_{ij} = b_j \right\}$$

Construyendo el Lagrangiano a partir del problema y sus respectivas restricciones:

$$\mathcal{L}(\pi, f, g) = \sum_{i,j} c_{ij} \pi_{ij} + \epsilon \sum_{i,j} \pi_{ij} (\log \pi_{ij} - 1) + \sum_i f_i \left(a_i - \sum_j \pi_{ij} \right) + \sum_j g_j \left(b_j - \sum_i \pi_{ij} \right).$$

Notar que:

$$\frac{\partial}{\partial \pi_{ij}} [\pi_{ij} (\log \pi_{ij} - 1)] = \log \pi_{ij}.$$

Derivando el Lagrangiano con respecto al plan de transporte:

$$\frac{\partial \mathcal{L}}{\partial \pi_{ij}} = c_{ij} + \epsilon \log \pi_{ij} - f_i - g_j.$$

Esto permite llegar a una condición de optimalidad que entrega información acerca de la estructura del plan de transporte π .

$$C_{ij} + \epsilon \log \pi_{ij} - f_i - g_j = 0$$

$$\log \pi_{ij} = \frac{f_i}{\epsilon} + \frac{g_j}{\epsilon} - \frac{C_{ij}}{\epsilon}$$

$$\pi_{ij} = e^{\frac{f_i}{\epsilon}} e^{\frac{g_j}{\epsilon}} e^{-\frac{C_{ij}}{\epsilon}}$$

Definiendo:

$$u_i = e^{\frac{f_i}{\epsilon}}, \quad v_j = e^{\frac{g_j}{\epsilon}}, \quad K_{ij} = e^{-\frac{C_{ij}}{\epsilon}}$$

El plan de transporte queda parametrizado elemento a elemento según:

$$\pi_{ij} = u_i K_{ij} v_j$$

De forma matricial:

$$\pi = \text{diag}(u) K \text{diag}(v)$$

Volviendo a la restricción sobre el primer marginal del plan de transporte π :

$$\sum_j \pi_{ij} = a_i \quad \Rightarrow \quad \pi \mathbf{1} = a.$$

Entonces:

$$\text{diag}(u) K \text{diag}(v) \mathbf{1} = a$$

$$\text{diag}(u) K v = a$$

Llegando a la expresión:

$$u = \frac{a}{Kv}$$

De forma similar, trabajando sobre el segundo marginal se llega a la expresión:

$$v = \frac{b}{K^t u}$$

Con estas dos expresiones para los vectores u , v y π se llega a un algoritmo iterativo que permite llegar a una solución numérica para la formulación de Kantorovich regularizada entrópicamente conocido como el algoritmo de Sinkhorn:

Definición 3.3.13 (Algoritmo de Sinkhorn). *Dadas dos medidas de probabilidad discretas μ y ν , una función costo c y un factor de regularización entrópica ϵ . Con K definido por*

$$K = e^{-\frac{c}{\epsilon}} \quad (3.36)$$

Construir vectores u^{k+1} y v^{k+1} :

$$u^{(k+1)} = \frac{a}{Kv^{(k)}} \quad v^{(k+1)} = \frac{b}{K^t u^{(k+1)}} \quad (3.37)$$

Calcular plan de transporte para $k + 1$ iteración:

$$\pi^{(k+1)} = \text{diag}(u^{(k+1)}) K \text{diag}(v^{(k+1)}) \quad (3.38)$$

Verificar que $\pi^{(k+1)}$ cumpla restricciones sobre sus marginales, de lo contrario iterar.

$$\pi^{(k+1)} \mathbf{1} = a \quad \pi^{T(k+1)} \mathbf{1} = b \quad (3.39)$$

3.3.6 Transporte óptimo desbalanceado

Si bien la formulación de Kantorovich con regularización entrópica es la versión mas relajada del problema de transporte óptimo, esta sigue teniendo una debilidad heredada desde la formulación de Monge. Esta debilidad corresponde a la conservación de masa entre distribuciones impuestas por 3.19 o 3.22, la cual obliga a que ambas distribuciones input posean "masas totales" de igual valor, lo cual se cumple de forma consistente trabajando con distribuciones de probabilidad. Esta limitación motiva la formulación del transporte óptimo desbalanceado (UOT, de Unbalanced Optimal Transport en inglés), presentado formalmente por primera vez en "Unbalanced Optimal Transport: Dynamic and Kantorovich Formulations" [6]. La idea central de UOT es reemplazar las estrictas restricciones que se tiene sobre los marginales por términos de penalización que miden la discrepancia entre los marginales del plan de transporte π y las distribuciones originales. De esta forma, el plan π ya no esta obligado a pertenecer al conjunto de planes de transporte admisibles $\Pi(\mu, \nu)$, si no que puede transgredir las restricciones marginales pagando un costo adicional proporcional, el cual esta controlado por un parámetro $\lambda > 0$.

Los terminos de penalización que utiliza la formulación desbalanceada del problema de transporte óptimo son expresadas mediante una medida matemática proveniente de la teoria de la información llamada divergencia de Kullback-Leibler.

Definición 3.3.14 (Divergencia de Kullback-Leibler). *Dadas dos distribuciones discretas P y Q , la divergencia de Kullback-Leibler de P con respecto a Q esta definida por:*

$$KL(P||Q) = \sum_i p_i \log\left(\frac{p_i}{q_i}\right) \quad (3.40)$$

La divergencia de Kullback-Leibler permite cuantizar el desempeño de la distribución candidato Q como representación de la distribución original P .

La divergencia de Kullback-Leibler satisface las siguientes propiedades que justifican su uso como término de penalización:

1. **No negatividad:** $KL(P||Q) \geq 0$, con igualdad si y solo si $P = Q$.
2. **Asimetría:** En general, $KL(P||Q) \neq KL(Q||P)$, por lo que no constituye una métrica en sentido estricto.
3. **Singularidad:** Si existe algún evento donde Q asigna probabilidad cero pero P no, entonces $KL(P||Q) = +\infty$.

Esta última propiedad es particularmente relevante en el contexto del transporte óptimo desbalanceado: una penalización infinita ocurre precisamente cuando los marginales del plan π se alejan completamente de las distribuciones originales, lo que dota a KL de un comportamiento naturalmente restrictivo.

Con estos elementos, es posible plantear formalmente el problema de transporte óptimo desbalanceado.

Definición 3.3.15 (Transporte Óptimo Desbalanceado). *Sean $\mu, \nu \in \mathcal{P}(X), \mathcal{P}(Y)$. Dado un parámetro de penalización $\rho > 0$, el problema de transporte óptimo desbalanceado consiste en encontrar el plan de transporte $\pi \in \Pi(\mu, \nu)$ que resuelve*

$$\min_{\pi \in \Pi(\mu, \nu)} \int_{X \times Y} c(x, y) d\pi(x, y) + \rho KL(\pi_1||\mu) + \rho KL(\pi_2||\nu) \quad (3.41)$$

donde π_1 y π_2 denotan las marginales del plan π sobre X e Y .

En esta formulación, los términos $KL(\pi_1||\mu)$ y $KL(\pi_2||\nu)$ penalizan la discrepancia entre las marginales del plan π y las distribuciones originales μ y ν respectivamente. El parámetro $\rho > 0$ controla el balance entre el costo de transporte y la penalización impuesta:

- Cuando $\rho \rightarrow \infty$, las penalizaciones se vuelven prohibitivamente costosas, forzando que $\pi_1 = \mu$ y $\pi_2 = \nu$. En este límite, el problema (3.3.15) recupera exactamente la formulación de Kantorovich, con π restringido al conjunto admisible $\Pi(\mu, \nu)$.

- Cuando $\rho \rightarrow 0$, las penalizaciones desaparecen y el plan óptimo tiende a concentrar toda la masa en la pareja de puntos (x, y) con menor costo $c(x, y)$, ignorando casi por completo las distribuciones originales.
- Para valores intermedios de ρ , el plan π puede transgredir las restricciones marginales pagando un costo proporcional, lo que permite trabajar con distribuciones cuyas masas totales difieren.

El problema de transporte óptimo desbalanceado también admite una formulación regularizada entrópicamente, cuya versión discreta se presenta a continuación.

Definición 3.3.16. *Formulación discreta del problema de transporte óptimo desbalanceado con regularización entrópica:*

Dado $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, factor de regularización entrópica ϵ y factor de penalización ρ .

$$\min_{\pi \in \Pi(\mu, \nu)} \left\{ \sum_{i,j} c_{ij} \pi_{ij} - \epsilon H(\pi) + \rho KL(\pi_1 \| \mu) + \rho KL(\pi_2 \| \nu) \right\} \quad (3.42)$$

Esta nueva formulación puede ser resuelta de forma computacional mediante una nueva versión del algoritmo de Sinkhorn donde los vectores u y v pasan a actualizarse de la siguiente manera:

$$u^{(k+1)} = \left(\frac{a}{K v^{(k)}} \right)^{\frac{\rho}{\rho+\epsilon}} \quad v^{(k+1)} = \left(\frac{b}{K^t u^{(k+1)}} \right)^{\frac{\rho}{\rho+\epsilon}} \quad (3.43)$$

Capítulo 4

Metodología

Tras haber establecido los fundamentos asociados al proceso de construcción de imágenes tipo B-mode en ultrasonido en el capítulo [3.2], así como las distintas formulaciones matemáticas del problema de transporte óptimo junto con sus respectivos métodos de resolución en el capítulo [3.3], el presente capítulo expone las metodologías desarrolladas con el propósito de evaluar el desempeño de técnicas de interpolación basadas en transporte óptimo aplicadas en distintas etapas del proceso de construcción de imágenes tipo B-mode.

La interpolación por transporte óptimo, al estar basada en un modelo matemático el cual compara distribuciones de probabilidad desde un punto de vista energético, puede interpretarse como una evolución geométrica entre dos distribuciones. A diferencia de métodos clásicos de interpolación, esta evolución tiende a capturar desplazamientos físicos y preservar estructuras de manera más consistente, evitando los efectos de difuminación asociados a interpolaciones lineales. Estas propiedades resultan particularmente interesantes en el contexto de la generación de imágenes ultrasonográficas, donde gran parte de la información presente en las señales de ultrasonido corresponden a ecos que sufren desplazamientos espaciales y redistribuciones de energía producto de la interacción de las ondas ultrasónicas con el tejido corporal.

En el presente trabajo se aplican técnicas de transporte óptimo sobre distintas representaciones de la señal de ultrasonido. Particularmente, se realizaron interpolaciones entre señales RF consecutivas pertenecientes a una misma adquisición, sobre las envolventes asociadas a las señales IQ beamformeadas, y además entre imágenes tipo B-mode correspondientes a adquisiciones consecutivas.

4.1 Interpolación sobre señal RF

Dado que las formulaciones de transporte óptimo discreto exigen, por lo mínimo, que ambas distribuciones sean completamente positivas, resulta inviable aplicar una interpolación baricéntrica

basada en transporte óptimo directamente a las señales RF crudas. Con el propósito de abordar esta limitación se exploró un enfoque basado en la interpolación de señales de audio mediante sus representaciones espacio-frecuenciales [7] y otro basado en interpolación por distancias TL^p , propuesta originalmente para el análisis de señales [8].

4.1.1 Interpolación entre espectrogramas

La idea central de este enfoque consiste en representar ambas señales mediante la Transformada de Fourier de Tiempo Corto (STFT), obteniendo así una representación tiempo-frecuencial compleja de cada una. A partir de esta representación, se extraen los espectrogramas de magnitud de ambas señales, donde cada ventana temporal define un espectro discreto de energía sobre el eje frecuencial.

Para cada instante de tiempo, los espectros de magnitud correspondientes en ambos espectrogramas son tratados como distribuciones discretas de probabilidad, lo que permite formular un problema de Transporte Óptimo entre ellas. La solución de este problema entrega un plan de transporte que describe como, para un mismo instante de tiempo, la energía espectral de la primera señal debe reorganizarse, de forma energéticamente óptima, para llegar a la distribución espectral de la segunda señal. Interpolando a lo largo de este plan de transporte se obtiene una distribución frecuencial intermedia para dicho instante.

Si se aplica este procedimiento para todos los instantes de tiempo en ambos espectrogramas, se construye un nuevo espectrograma de magnitud interpolado. Dado el carácter complejo de la STFT, se requiere además reconstruir una fase coherente para el espectrograma, lo cual se realiza mediante un mecanismo tipo "phase vocoder", siguiendo el esquema propuesto en [7]. Finalmente, la señal RF interpolada es obtenida aplicando Transformada Inversa de Fourier de Tiempo Corto (ISTFT) sobre el espectrograma de magnitud interpolado junto con su respectiva fase reconstruida.

El primer paso para implementar esta metodología es crear dos funciones. Una que llegue a un plan de transporte óptimo mediante iteraciones de Sinkhorn y otra que interpole a lo largo de dicho plan de transporte.

Las funciones descritas en A.1 y A.2 trabajan en conjunto para construir una distribución discreta de probabilidad *interpolada* entre otras dos a y b . Como las magnitudes frecuenciales que la STFT entrega para cada instante de tiempo no son distribuciones de probabilidad, antes de aplicar técnicas de interpolación por OT se debe aplicar un proceso de normalización y posterior escalamiento para obtener una nueva distribución de energía espectral. Es por esto que se construye una función auxiliar que entregue el output usual de la STFT junto con las magnitudes espectrales para cada instante normalizadas, acompañadas de sus respectivas normas para así reescalar después de la interpolación.

Habiendo construido esta función auxiliar A.3 junto con las funciones que aplican interpolación de OT entre dos distribuciones de probabilidad, se procede a construir nuevas funciones que

confeccionen una fase coherente para complementar la magnitud espectral interpolada.

Habiendo construido todas las funciones necesarias para realizar una interpolación basada en OT entre dos espectrogramas de magnitud de señales distintas y funciones que construyen una fase coherente para dicho espectrograma de magnitud interpolado en A.4, se procede a construir la función principal A.5 que permitirá interpolar una señal RF artificial entre otras dos señales RF reales.

Para dejar en claro de mejor manera la forma en la que opera la función principal se presenta un ejemplo en el cual se muestra los procesos por los cuales un par de señales RF pasan para interpolar otra RF artificial. Considerar las dos señales $s1$ y $s2$ representadas por series de tiempo de amplitud con 1600 muestras cada una:

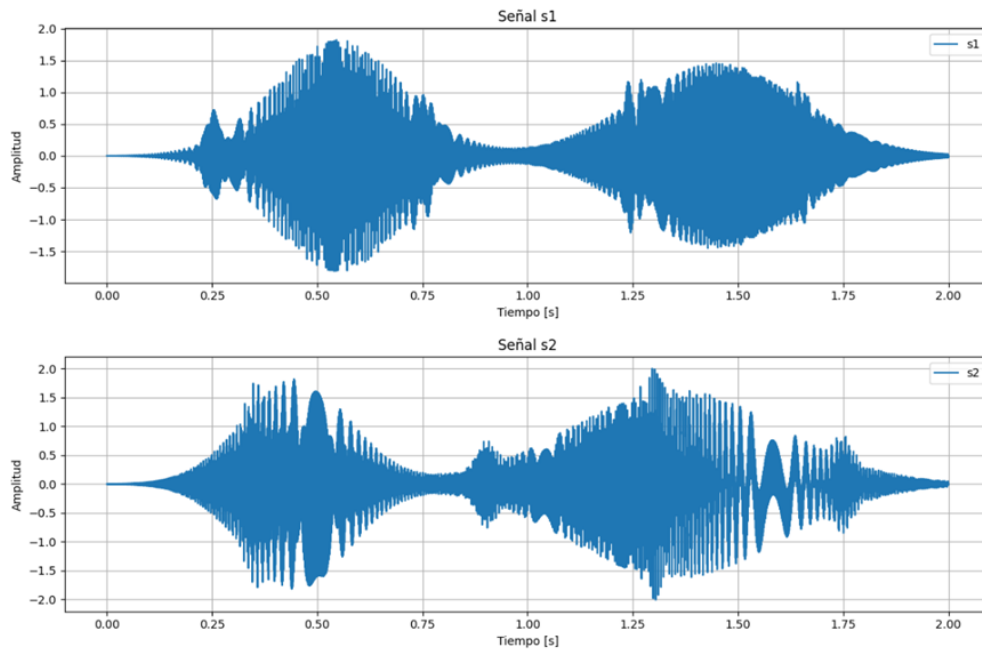


Figura 4.1: Señales 1 y 2 para ejemplificar el metodo de interpolación OT sobre espectrogramas

Se procede a calcular los espectrogramas de magnitud de ambas señales mediante STFT. Para efectos de este ejemplo se utilizó ventanas de Hann de 512 muestras cada una, con un solapamiento de 384 entre ventanas consecutivas.

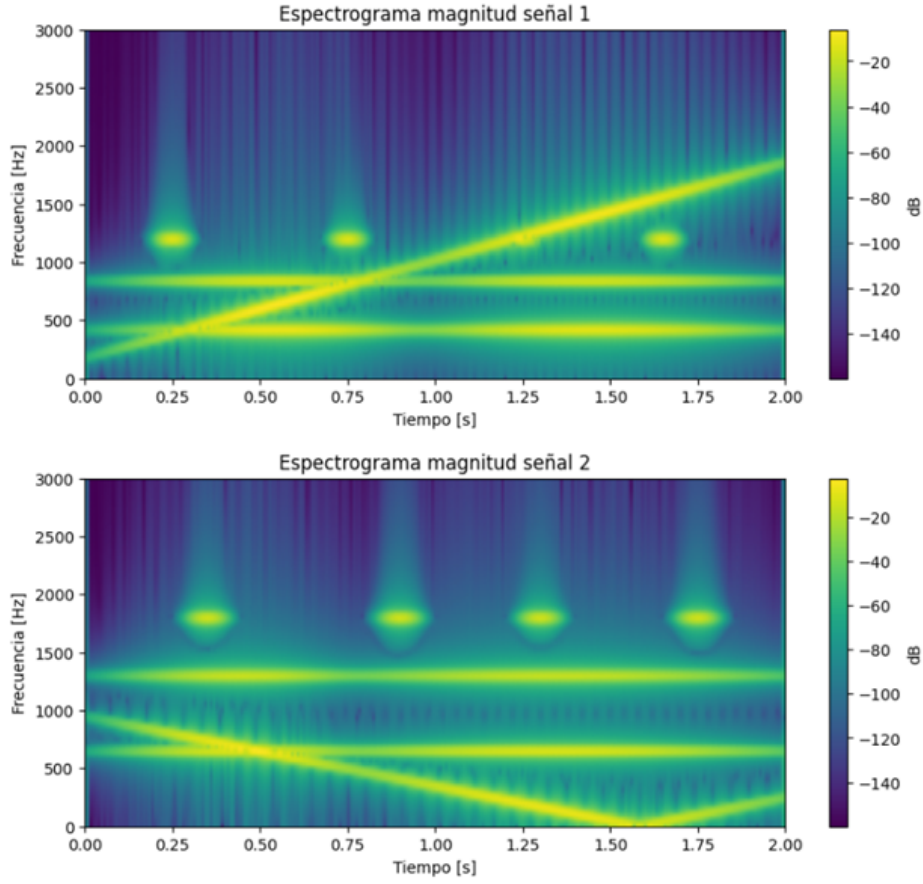


Figura 4.2: Espectrogramas de magnitud de las señales 1 y 2.

La interpolación OT ocurre entre los arreglos que representan la energía espectral para un instante determinado en ambas señales. Es decir, se estaría considerando cada columna de ambos espectrogramas como distribuciones de probabilidad y la interpolación ocurre entre las columnas correspondientes al mismo instante de tiempo de ambos espectrogramas. Por ejemplo, para $t = 0$ las distribuciones a y b entre las que se interpola vendrían siendo las primeras columnas de ambos espectros pero normalizadas para que sumen 1:

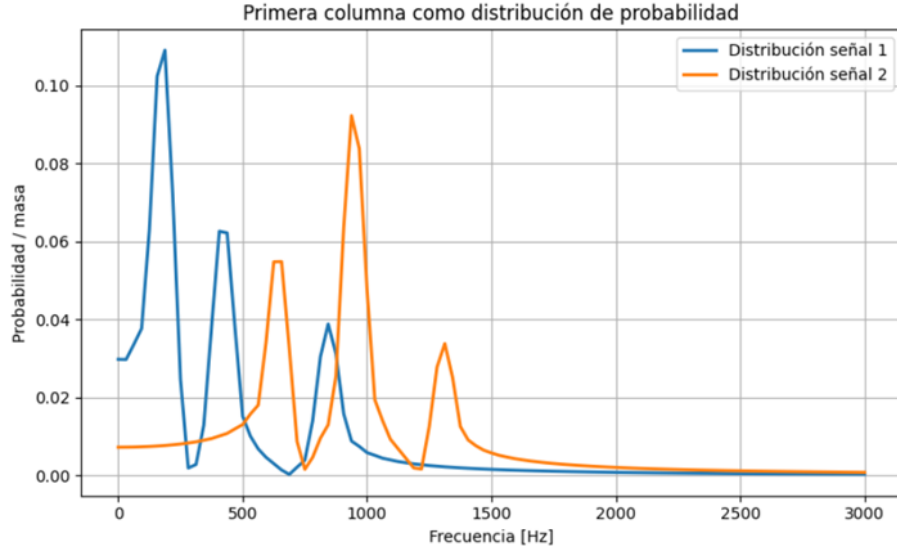


Figura 4.3: Distribuciones de probabilidad correspondientes a la energía espectral de s_1 y s_2 para $t = 0$.

Una vez elegidas el par de distribuciones entre las que se interpola se procede a resolver el problema de transporte óptimo con regularización entrópica correspondiente, obteniendo así un plan de transporte con el cual se interpola una nueva distribución con un factor α determinado, por ejemplo $\alpha = 0.5$ entrega:

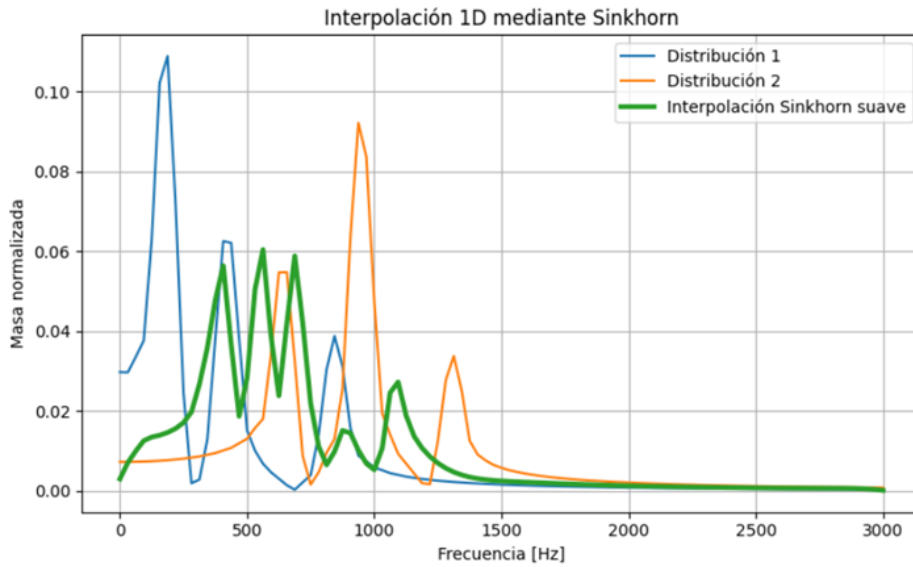


Figura 4.4: Distribución de probabilidad interpolada con factor $\alpha = 0.5$ entre columnas frecuenciales de espectrogramas de s_1 y s_2 para $t = 0$

Una vez realizado este proceso para todos los instantes de tiempo comprendidos en los espectrogramas de ambas señales se llega a un espectrograma de magnitud interpolado mediante

técnicas OT.

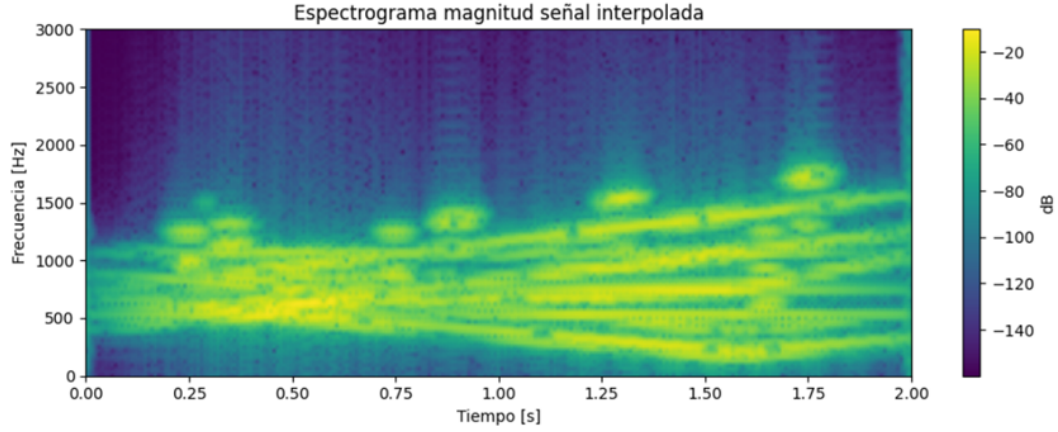


Figura 4.5: Espectro de magnitud interpolado mediante técnicas OT entre las señales s_1 y s_2

Con la información de este espectrograma de magnitud interpolado se reconstruye una fase coherente, guiado por el procedimiento de phase-vocoder propuesto por [7]. Al aplicar ISTFT con los mismos parámetros que se aplicaron sobre las señales s_1 y s_2 en un principio, finalmente se llega a una señal interpolada.

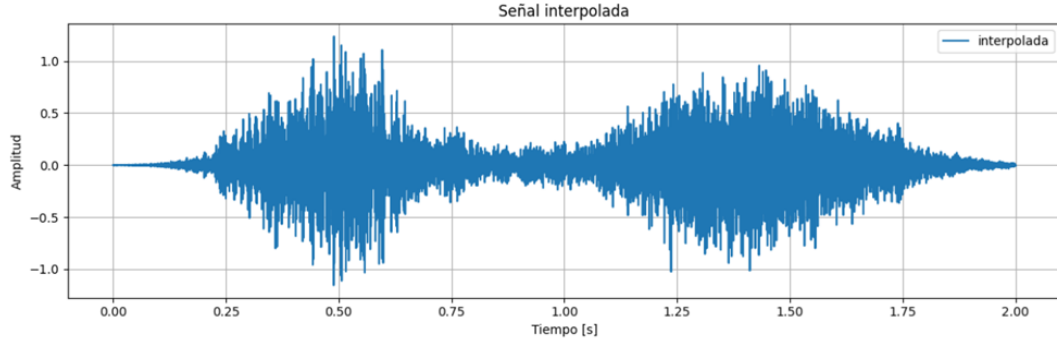


Figura 4.6: Señal interpolada entre s_1 y s_2 mediante OT en espectros de magnitud

4.1.2 Interpolación distancias TL^p

Con el propósito de abordar la limitación de estricta positividad asociada a las formulaciones clásicas de transporte óptimo, se exploró una estrategia de interpolación basada en las distancias TL^p propuestas en [8]. Estas corresponden a una generalización del transporte óptimo diseñada para comparar señales generales sin requerir condiciones estrictas de positividad o normalización.

A diferencia de las distancias de Wasserstein clásicas, las métricas TL^p incorporan simultáneamente información geométrica y diferencias de amplitud dentro del funcional de costo, permitiendo trabajar directamente sobre señales oscilatorias como las señales RF de ultrasonido.

La idea principal de la métrica TL^p consiste en representar cada punto espacial $x \in X$ asociado a una señal $f : X \rightarrow \mathbb{R}$ mediante el par $(x, f(x))$, construyendo así una representación conjunta posición-amplitud de la señal, la cual será utilizada para construir el funcional de costo. Por consiguiente, la información relevante de la señal es incorporada principalmente dentro del funcional de costo, mientras que las medidas espaciales μ y ν , las cuales previamente representaban cuánta masa era asignada a cada posición del dominio, pasan a representar únicamente la importancia o peso relativo asociado a cada posición espacial de la señal.

Dadas dos funciones $f, g : X \rightarrow \mathbb{R}$, la distancia TL^p se define como

$$d_{TL^p}^p(f, g) = \min_{\pi \in \Pi(\mu, \nu)} \int_{X \times X} \left(\frac{1}{\lambda} |x - y|^p + |f(x) - g(y)|^p \right) d\pi(x, y) \quad (4.1)$$

donde π corresponde a un plan de transporte admisible entre las medidas espaciales asociadas a ambas señales μ y ν , $\lambda > 0$ controla el compromiso entre el costo geométrico y el costo asociado a diferencias de amplitud.

Bajo esta formulación, la distancia no solo penaliza diferencias en amplitud entre señales, sino también desplazamientos espaciales entre sus estructuras características, permitiendo modelar simultáneamente deformaciones y variaciones locales de intensidad.

A continuación, en A.6 se presenta un código PYTHON el cual interpola entre dos señales mediante distancias TL^p :

Dado que la formulación TL^p corresponde a una variante del problema de transporte óptimo con un funcional de costo específico, el problema puede resolverse mediante iteraciones de Sinkhorn de manera análoga a las formulaciones regularizadas de transporte óptimo clásico.

En la implementación utilizada en este trabajo, las medidas espaciales discretas a y b fueron definidas como distribuciones uniformes, asignando la misma importancia relativa a cada posición temporal de las señales. Adicionalmente, previo a la construcción de la matriz de costo, las amplitudes de ambas señales fueron normalizadas respecto al máximo valor absoluto presente en las señales analizadas. Esto permitió mantener en órdenes de magnitud comparables tanto el término geométrico como el término asociado a amplitud dentro del funcional de costo, evitando que uno de ellos dominara numéricamente el problema de transporte. Finalmente, una vez obtenida la señal interpolada, esta fue reescalada utilizando el mismo factor de normalización para recuperar la amplitud original de las señales de entrada.

Dado que las señales RF utilizadas en ultrasonido suelen contener una gran cantidad de muestras, aplicar directamente la formulación TL^p sobre señales completas implica construir matrices de costo de tamaño $N \times N$, lo cual produce un costo computacional y de memoria del orden $\mathcal{O}(N^2)$. Bajo estas condiciones, la interpolación directa de señales RF completas resulta computacionalmente poco eficiente e incluso impracticable para señales de gran longitud. Con el propósito de reducir esta

complejidad computacional, se optó por una estrategia de interpolación por bloques. Para ello, cada señal fue dividida en segmentos de menor tamaño sobre los cuales se aplicó de manera independiente el procedimiento de interpolación basado en TL^p . Posteriormente, las señales interpoladas obtenidas para cada bloque fueron reconstruidas mediante una estrategia overlap-add (OLA), permitiendo combinar coherentemente las contribuciones locales de cada segmento y reducir discontinuidades entre bloques consecutivos. La función que realiza la reconstrucción OLA se muestra en A.7.

4.1.3 Evaluación de interpolaciones OT sobre señales RF

La forma en que se evaluó el desempeño de las señales RF interpoladas consistió en analizar el impacto que estas tienen sobre la reconstrucción de imágenes B-mode. Para ello, se utilizaron adquisiciones obtenidas mediante un transductor lineal de 128 elementos, donde cada elemento proporciona una señal RF independiente correspondiente a un canal de recepción.

Con el objetivo de evaluar la capacidad de los distintos métodos para reconstruir información faltante, se eliminaron canales intermedios dentro del arreglo de señales RF y posteriormente se estimaron mediante interpolación. Las señales faltantes fueron reconstruidas utilizando tres enfoques diferentes: interpolación lineal, interpolación basada en Transporte Óptimo (OT) e interpolación basada en Transporte Óptimo con métrica TL^p .

Una vez obtenidas las señales interpoladas, estas fueron incorporadas nuevamente al conjunto de datos RF para la reconstrucción de las correspondientes imágenes B-mode. El desempeño de cada método fue evaluado comparando las imágenes generadas a partir de los datos interpolados con aquellas obtenidas utilizando la adquisición original completa. De esta manera, la calidad de la interpolación se cuantificó de forma indirecta a través de su efecto sobre la formación de imagen ecográfica, permitiendo analizar la capacidad de cada método para preservar la información espacial y estructural presente en los datos originales.

4.2 Interpolación sobre envolventes post-beamforming

Además de las estrategias de interpolación aplicadas directamente sobre señales RF, se exploró una metodología basada en la interpolación de envolventes post-beamforming. En este caso, el proceso de formación de imagen fue realizado previamente mediante beamforming DAS sobre señales IQ, obteniendo posteriormente las envolventes que conforman la representación tipo B-mode.

Bajo esta formulación, la interpolación ya no se realiza sobre señales RF oscilatorias, sino directamente sobre los valores de intensidad que componen cada píxel de la imagen ultrasonográfica. Dado que las envolventes corresponden a magnitudes positivas asociadas a la energía local de la señal ultrasónica, estas pueden interpretarse naturalmente como distribuciones discretas no negativas,

permitiendo aplicar directamente formulaciones clásicas de transporte óptimo y transporte óptimo desbalanceado.

La principal motivación de esta aproximación consiste en evaluar el comportamiento del transporte óptimo sobre representaciones donde gran parte de la información de fase ya ha sido eliminada mediante la detección de envolvente, reduciendo así las oscilaciones rápidas presentes en las señales RF originales y facilitando la construcción de distribuciones más estables para el problema de transporte. Para ello se construirán funciones que realicen un proceso de "upsampling" lateral sobre imágenes B-mode ya confeccionadas, es decir, agregar columnas interpoladas mediante distintos métodos para observar el comportamiento de los métodos de interpolación.

Durante esta y posteriores secciones se trabaja en MATLAB dada la existencia de una toolbox de Ultrasonido (MUST), la cual facilita bastante el proceso de construcción de imágenes B-mode. Habiendo entregado esta información, el primer paso fue construir funciones MATLAB que interpolen una distribución discreta mediante UOT y mediante la solución numérica de OT 1D mediante CDFs.

La función A.8 resuelve el problema de transporte óptimo desbalanceado entre $col1$ y $col2$ mediante iteraciones de Sinkhorn con parámetro de regularización ϵ y parámetro de desbalanceo ρ , obteniendo un plan de transporte P con el cual se calcula una interpolación baricéntrica col_{mid} de factor α .

La función A.9 llega a una distribución interpolada col_{mid} entre $col1$ y $col2$ mediante la solución analítica de transporte óptimo balanceado 1D mediante CDFs. Dada esta solución no es necesario calcular de forma explícita el plan de transporte P .

Como se mencionó anteriormente, las envolventes post-beamforming comprenden un arreglo de intensidades sonoras asignadas a grilla de píxeles, la cual al pasar por un proceso de compresión logarítmica se transforma en una imagen B-mode. Es por esto que a continuación se construyen funciones que toman arreglos de señales envolventes e interpolan entre cada par de columnas otra adicional mediante UOT, OT 1D. También se construye una función que realiza un proceso de upsampling lineal para comparar.

La función descrita en A.10 toma un arreglo de señales envolventes de dimensiones $[N_z, N_x]$, donde N_z vendría siendo el número de píxeles de profundidad y N_x el número de píxeles en la dirección lateral. Dependiendo del número de valores de α distintos que se entreguen, la función interpolará N_α columnas adicionales entre cada par de columnas. Para construir una función que realice esta misma tarea pero con un upsampling basado OT 1D mediante CDFs basta con reemplazar la función `uot_inter_col` de la línea número 19 por `ot_cdf_interp_col`.

4.3 Interpolación sobre matrices B-mode

Hasta este momento, en este documento solo se ha trabajado sobre señales ultrasónicas que pueden ser interpretadas como distribuciones unidimensionales. Sin embargo, las imágenes B-mode corresponden a distribuciones definidas sobre un dominio bidimensional, donde cada píxel posee una intensidad asociada a una posición espacial dada por sus coordenadas de profundidad y dirección lateral.

Una forma conveniente de abordar este problema consiste en representar la imagen bidimensional como una distribución discreta sobre una grilla de tamaño $(N_z \times N_x)$. Posteriormente, dicha distribución puede vectorizarse, transformando la matriz de intensidades en un vector de longitud $(N_z N_x)$. De esta manera, el algoritmo de Sinkhorn puede aplicarse directamente sobre las distribuciones vectorizadas, sin necesidad de modificar su formulación matemática.

No obstante, la vectorización por sí sola no preserva la estructura espacial de la imagen. Para mantener la geometría bidimensional original es necesario definir una función de costo que incorpore la posición espacial de cada píxel. Para ello, a cada elemento del vector se le asocian las coordenadas $((z_i, x_i))$ del píxel correspondiente, y el costo de transporte entre dos píxeles i y j se define mediante la distancia euclidiana cuadrática $c_{ij} = |(z_i, x_i) - (z_j, x_j)|^2$. De esta forma, aunque las distribuciones sean manipuladas computacionalmente como vectores unidimensionales, la información geométrica bidimensional de la imagen permanece incorporada en la matriz de costos utilizada por el algoritmo de Sinkhorn. Como consecuencia, el transporte calculado refleja desplazamientos espaciales reales dentro de la imagen B-mode y no simplemente reordenamientos arbitrarios de intensidad.

En A.12 se muestra la función MATLAB que toma un par de arreglos bidimensionales P_1 y P_2 entre los cuales interpola otro *Barys* mediante técnicas basadas en transporte óptimo. El primer paso de la función presentada consiste en almacenar la masa total de los arreglos bidimensionales que serán interpolados. Para ello, se definen las variables M_1 y M_2 , dadas por la suma de todos los elementos de P_1 y P_2 , respectivamente. Posteriormente, ambos arreglos son vectorizados por columnas, obteniéndose los vectores a y b , los cuales representan las distribuciones origen y destino sobre una grilla discreta de $N_z \times N_x$ puntos.

Dado que el algoritmo de Sinkhorn opera sobre distribuciones discretas representadas como vectores, esta vectorización permite aplicar directamente la formulación de transporte óptimo regularizado. Sin embargo, para no perder la información espacial propia de la imagen bidimensional, se construye una matriz de costo a partir de las coordenadas originales de cada píxel. Así, a cada elemento vectorizado se le asocia una posición $\mathbf{r}_i = (z_i, x_i)$, donde z_i corresponde a la coordenada en profundidad y x_i a la coordenada lateral. La matriz de costo se define entonces como $C_{ij} = |\mathbf{r}_i - \mathbf{r}_j|^2$, lo que corresponde a la distancia euclidiana cuadrática entre el píxel de origen i y el píxel de destino j . De esta manera, aunque las imágenes sean manipuladas computacionalmente como

vectores, la geometría bidimensional se conserva mediante la función costo. Esto permite que el plan de transporte P calculado por Sinkhorn penalice desplazamientos espaciales grandes y favorezca movimientos coherentes dentro del dominio de la imagen.

El proceso de búsqueda de plan de transporte entre las distribuciones a y b termina en la línea 33 de A.12 llegando a una matriz P , cuyos elementos P_{ij} indican cuanta masa se mueve desde la posición i en a a la posición j en b , arreglos que representan las imágenes originales P_1 y P_2 en forma vectorizada. Con este plan de transporte P se construyen las imágenes interpoladas para cada valor de α entregado. Para ello, en las líneas 50 y 51 se calcula la posición intermedia asociada a cada elemento P_{ij} , obteniéndose las coordenadas donde deberá depositarse la masa transportada entre los píxeles de origen y destino. Dado que estas posiciones intermedias no necesariamente coinciden con índices enteros de la grilla de la imagen, la masa asociada a cada elemento P_{ij} no es asignada a un único píxel. En su lugar, ésta se distribuye entre los cuatro píxeles vecinos que rodean la posición interpolada mediante una interpolación bilineal, permitiendo una representación espacial más suave y preservando la masa total transportada.

Al igual que en las interpolaciones basadas en distancias TL^p , la aplicación directa del algoritmo descrito en A.12 sobre imágenes B-mode completas presenta un costo computacional elevado. En efecto, para una imagen de dimensiones $N_z \times N_x$, la matriz de costo posee dimensiones $N_z N_x \times N_z N_x$, por lo que el tiempo de cálculo aumenta rápidamente en función del tamaño de las imágenes entre las que se interpola.

De forma análoga a lo realizado en la interpolación de señales RF mediante bloques y reconstrucción overlap-add, las imágenes B-mode fueron divididas en ventanas superpuestas para limitar el tamaño de los problemas de transporte óptimo resueltos individualmente.

Se observa en A.13 cada par de bloques interpolados mediante Sinkhorn es ponderado por una ventana bidimensional de Hann y acumulado en la posición espacial correspondiente. Debido al solapamiento entre bloques, un mismo píxel puede recibir contribuciones desde múltiples ventanas. Para compensar este efecto, se acumulan simultáneamente los pesos de las ventanas y posteriormente se realiza una normalización punto a punto. Este procedimiento corresponde a una reconstrucción tipo Constant Overlap-Add (COLA), permitiendo obtener una imagen final continua y libre de discontinuidades visibles entre bloques adyacentes.

4.4 Evaluación de interpolaciones

La evaluación de las interpolaciones basadas en técnicas OT en distintas etapas del proceso de generación de imágenes ultrasónicas tipo B-mode se realizó principalmente mediante la comparación de resultados con algún dato real disponible o "ground truth". A continuación se presenta como se evalúan los resultados obtenidos en cada etapa de la generación de imagen.

4.4.1 Interpolación sobre señales RF

Dado que la imagen B-mode se obtiene a partir de un arreglo de señales RF correspondiente a los ecos de retorno recibidos por cada uno de los elementos piezo-electricos del transductor, el dato "ground truth" con el que se compara son las señales RF originales junto con la respectiva imagen B-mode que ellas producen.

Los metodos de interpolación OT se aplicarán sobre parejas de señales RF pertenecientes a la adquisición de datos original. Se tiende a pensar en interpolar las señales entre canales consecutivos, sin embargo al hacer esto se pierde la posibilidad de comparár el resultado con un "ground truth", ya que no se posee una señal RF correspondiente a una posicion intermedia entre dos elementos piezo-electricos. Es por esto que se propone interpolar entre señales RF de forma intercalada para así tener un dato real con el cual comparar.

Para construir los casos de interpolación se utilizaron pares de canales separados por una posición lateral. En particular, los canales con índices pares fueron utilizados como señales de referencia, mientras que los canales impares fueron retirados y utilizados posteriormente como *ground truth*. De esta forma, las parejas de señales RF (0, 2), (2, 4), (4, 6), ... permitieron estimar respectivamente los canales intermedios 1, 3, 5, ... mediante los distintos métodos de interpolación evaluados.

Dado que se trabajó con un transductor lineal L11-5v de 128 canales, con el fin de mantener un esquema homogéneo de interpolación no se alteraron los datos correspondientes a los elementos 63 y 64 para evitar casos particulares en los extremos del arreglo en los cuales no se posee señales vecinas suficientes para formar una pareja de señales RF entre las cuales interpolar. A continuación se muestra una figura que muestra los canales sobre los cuales se interpola mediante metodos OT.



Figura 4.7: Elementos de transductor para evaluación de interpolación sobre señales RF. En azul se encuentran los elementos sin alterar, mientras que en rojo los elementos sobre los cuales se sobreescribe una interpolación

Con esta metodología se evaluará el desempeño de técnicas de interpolación basadas en OT entre parejas de elementos azules para obtener una señal RF intermedia que sobreescribirá la señal original de cada elemento rojo. Esta nueva señal RF interpolada por OT se comparará con su respectiva señal RF original correspondiente a su posición, tambien se comparará su desempeño respecto a una interpolación lineal. Tambien se evaluará de forma cualitativa el efecto que estas interpolaciones tienen sobre la imagen B-mode final.

4.4.2 Interpolación sobre envolventes post-beamforming

Dentro del proceso de construcción de imágenes B-mode, después del proceso de beamforming DAS se obtiene un arreglo de píxeles con intensidades acústicas asignadas, el cual después de pasar por una compresión logarítmica pasa a su estado final de imagen B-mode. Las funciones definidas en 4.2 consideran columnas de este arreglo de píxeles como distribuciones de probabilidad, permitiendo interpolar mediante OT columnas de píxeles adicionales.

De forma similar a como se evalúa la interpolación sobre señales RF, se retirarán columnas de píxeles en posiciones impares dentro del arreglo y se reemplazarán por interpolaciones basadas en OT y lineales. De esta forma se tendrá un dato "ground truth", correspondiente a la columna de píxeles original, con el cual comparar. Además se evaluará de forma cualitativa el efecto que estas interpolaciones tienen sobre la imagen B-mode final.

4.4.3 Interpolación sobre matrices B-mode

Las matrices B-mode son la representación "frame a frame" del video obtenido durante una adquisición de datos ultrasónicos, donde cada frame es una imagen B-mode de la región estudiada durante un instante particular. Las funciones definidas en 4.3 permiten interpolar entre parejas de frames para obtener una serie de frames interpolados a través del plan de transporte óptimo definido entre los dos frames originales.

Para evaluar el desempeño de la interpolación, dentro de la matriz B-mode se definirán parejas de frames entre los cuales se interpolará, para posteriormente reemplazar todos los frames intermedios por frames interpolados mediante técnicas OT y lineales, donde los datos "ground truth" corresponden a los frames B-mode originales correspondientes a los instantes temporales sobre los cuales se interpoló. Además se evaluará cualitativamente el efecto que estas técnicas de interpolación tienen sobre el video de una adquisición ultrasónica.

Capítulo 5

Resultados

Los datos ultrasónicos sobre los que se trabajó fueron obtenidos mediante adquisiciones de datos mediante un equipo "Verasonics Vantage" utilizando un transductor lineal "L11-5V" de 128 elementos piezo-electricos. El equipo computacional con el que se hizo el post-tratamiento de los datos correspondiente a las interpolaciones basadas en OT fue un laptop "Lenovo LOQ 15IAX9" equipado con un procesador "Intel Core i5" de doceava generación junto con 16 GB de RAM DDR5 de 4800 MHz.

5.1 Interpolación sobre señal RF

La adquisición de datos ultrasónicos sobre la cual se trabajó para interpolar señales RF mediante OT corresponden a la toma de datos de un fantoma con una esfera insertada en ella, la cual actua como reflectante de ondas acústicas. En la figura 5.1 se muestra la imagen modo B construida a partir de los datos RF sobre los cuales se trabaja.

En particular, el archivo MATLAB almacena las señales RF en un arreglo de dimensiones $[128, 2867200]$ donde la primera dimensión corresponde al numero de canales del transductor L11-5v y la segunda corresponde al numero de muestras que utilizan las series de tiempo de amplitud para representar cada señal RF. Este gran numero de muestras es el principal obstaculo al momento de aplicar interpolaciones tipo OT directamente sobre las señales RF, particularmente al momento de construir el plan de transporte π , el cual en este caso seria una matriz de dimensiones $[2867200, 2867200]$, ya sea para utilizarlo en una interpolación o simplemente para verificar si se cumplen las restricciones sobre sus marginales durante una iteración de Sinkhorn.

Antes de mostrar la imagen B-mode que los nuevos datos RF construidos al reemplazar cada señal en una posición roja de la figura 4.7 por una interpolada con técnicas OT, en la figura 5.2 se presenta una donde se reemplazan por señales interpoladas linealmente entre sus vecinas.

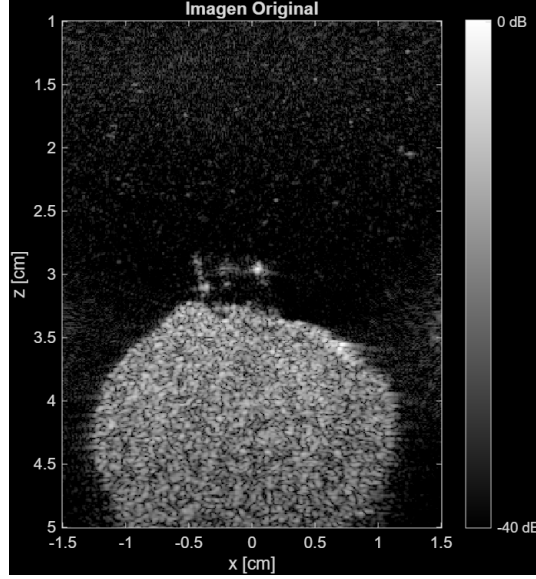


Figura 5.1: Imagen B-mode construida a partir de datos RF no intervenidos

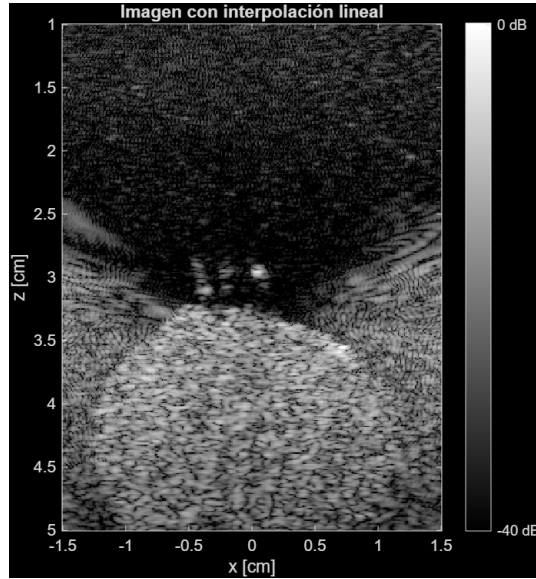


Figura 5.2: Imagen B-mode construida a partir de datos RF con señales RF interpoladas linealmente

5.1.1 Interpolación entre espectrogramas

Sabiendo que la adquisición de datos con la que se trabajó tiene una frecuencia de muestreo de $f_s = 31.25 \times 10^6$, para construir los espectrogramas de cada señal se considerarán los siguientes parametros para la transformada de Fourier de tiempo corto (STFT):

- Frecuencia de muestreo: $f_s = 31.25 \times 10^6$

- Función ventana: Hamming
- Numero de muestras por ventana: 1024
- Solapamiento entre ventanas: %50

Aplicar STFT con estos parametros permite obtener facilmente la representación tiempo-frecuencial compleja de cada señal, la cual al pasar por una función valor absoluto se convierte en un espectrograma de magnitud, permitiendo utilizar las funciones definidas en el la sección 4.1.1 del marco teórico para obtener un espectrograma de magnitud interpolado "columna a columna" mediante iteraciones de sinkhorn con un factor de regularización $\epsilon = 0.01$.

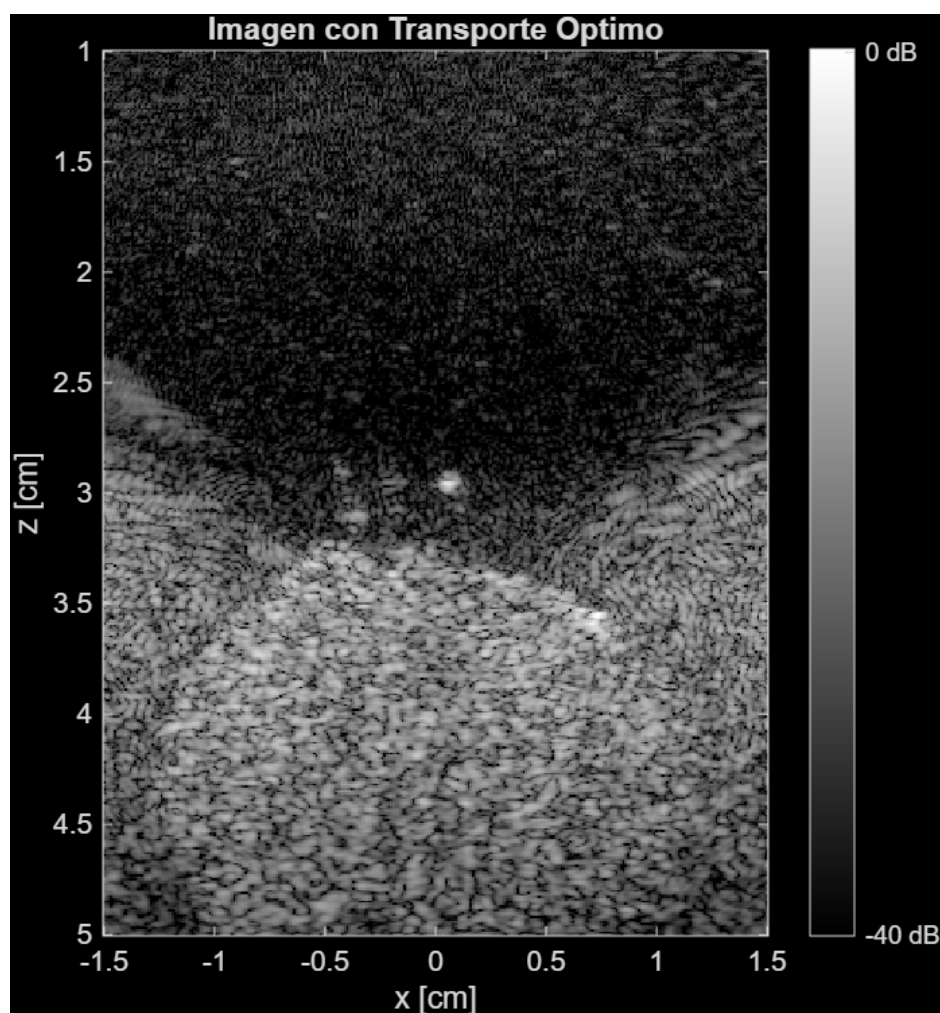


Figura 5.3: Imagen B-mode construida a partir de datos RF con señales RF interpoladas por OT regularizado

5.1.2 Interpolación con distancias TL^p

A diferencia de las formulaciones clásicas de OT, donde las distribuciones μ y ν representan la distribución espacial de una cantidad de masa en específico, las distancias TL^p interpretan estas medidas no como masa, si no como una medida de importancia relativa de cada muestra de la señal RF con respecto al resto de las demás, dejando toda la información característica de la señal en el funcional de costo el cual castiga diferencias temporales y de amplitud entre muestras.

Teniendo este detalle en cuenta, se consideró que ambas medidas μ y ν se consideran como arreglos uniformes de largo [286720], otorgando así la misma importancia a todas las muestra que componen las señales. Dentro del funcional de costo se utilizaron factores $\lambda = 0.05$ y $p = 2$ penalizando cuadráticamente el transporte entre muestras muy alejadas en el eje temporal. Para las iteraciones de Sinkhorn se utilizó un factor de regularización entrópica $\epsilon = 0.01$.

Para el proceso final de Overlap-Add sobre los bloques individuales interpolados mediante distancias TL^p se utilizaron los siguientes parámetros:

- Función ventana: Hamming
- Número de muestras por bloque/ventana: 512
- Solapamiento entre ventanas: %50

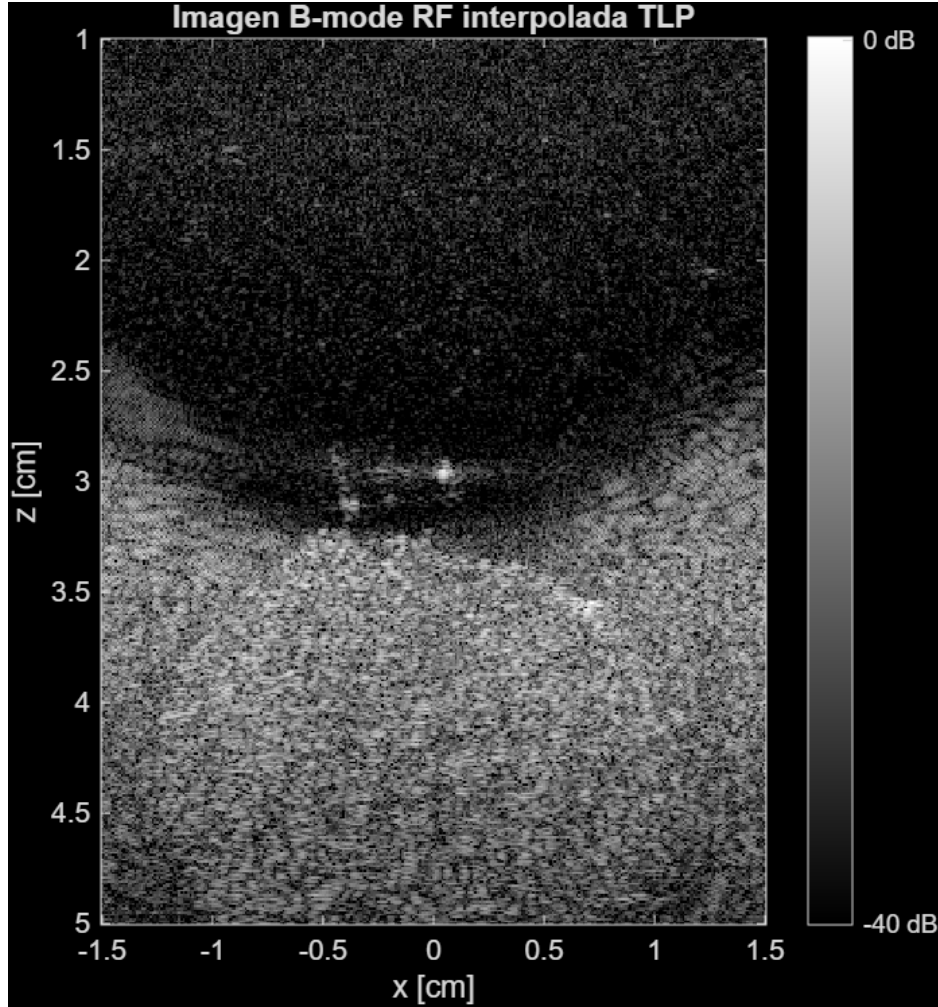


Figura 5.4: Imagen B-mode construida a partir de datos RF con señales RF interpoladas por OT con distancias TLP

5.1.3 Análisis de resultados

Al observar las imagenes B-mode de las señales RF intervenidas de distintas formas en las figuras 5.2, 5.3 y 5.4 es facil ver que al alterar en lo mas mínimo las señales RF originales se generan artefactos visuales no deseados en la imagen B-mode final. Esto sucede porque al interpolar una señal se destruyen pequeños detalles en su fase, los cuales originalmente daban lugar a interferencias constructivas y destructivas durante el proceso de beamforming. El beamforming DAS asume que las señales RF de todos los elementos del transductor corresponden a "una misma escena" y por tanto, mantienen una relación de coherencia física entre sí. Cuando parte de estas señales son reemplazadas por señales interpoladas que no preservan por completo dicha coherencia, se producen pérdidas de resolución, artefactos visuales y alteraciones en los patrones en la imagen B-mode final.

A pesar de los artefactos que produce aplicar cualquier tipo de interpolación sobre las señales RF se comparará cuantitativamente cada una de las señales interpoladas con su respectivo dato "ground truth" original mediante metricas RMSE y distancia de Wasserstein orden 2.

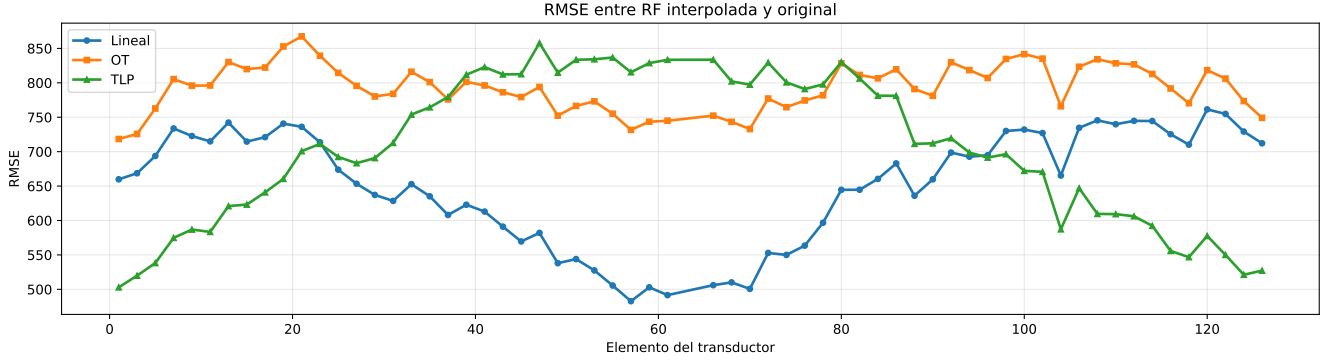


Figura 5.5: RMSE de señales RF interpoladas con respecto a señales RF originales

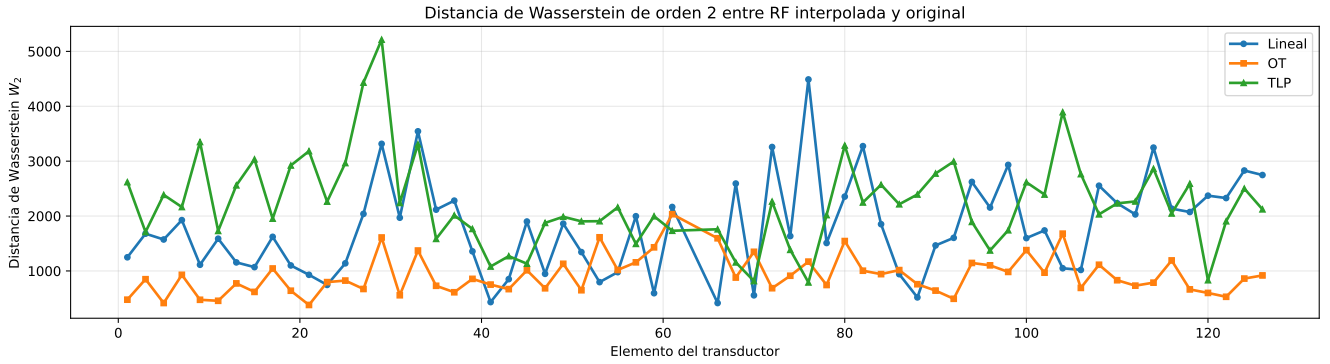


Figura 5.6: Distancia de Wasserstein orden 2 entre señales RF interpoladas y señales RF originales

Si bien el RMSE constituye una métrica ampliamente utilizada para cuantificar diferencia entre señales, esta presenta ciertas limitaciones en el contexto de comparar señales RF. Al tratarse de una comparación de amplitud entre dos señales para un mismo instante de tiempo, el RMSE evalúa exclusivamente el "solapamiento" entre las señales que se comparan. Por consecuencia, pequeños desplazamientos de los máximos locales en las señales pueden generar grandes incrementos en el error. Es por esto que se utiliza la métrica de Wasserstein orden 2 la cual en lugar de evaluar diferencias locales de amplitud, evalúa similitudes en la estructura geometrica entre dos señales.

En la figura 5.5 se observa que todos los metodos aplicados para interpolar señales RF presentan valores de RMSE muy elevados, especialmente los metodos basados en OT. Esto se da por consecuencia de la comparación punto a punto que la RMSE realiza, ademas de el gran largo de las señales RF que exacerba los valores de este error.

La figura 5.6 muestra la distancia de Wassertein orden 2 entre la señal interpolada y la original, evaluando que tan costoso es reordenar las series de tiempo de la señal RF interpolada para obtener la señal RF original. De esta figura se puede apreciar que el metodo de interpolación de espectros de magnitud mediante OT fue el que tuvo mejores resultados, es decir señales RF que presentaban una estructura geométrica de mayor similitud a la original. Sin embargo, como se vio en las figuras 5.2, 5.3 y 5.4, esta similitud no es suficiente para generar una imagen B-mode coherente.

5.2 Interpolación sobre envolventes post-beamforming

De igual manera a como se hizo con la interpolación directa sobre señales RF, para esta sección se presentará nuevamente la imagen B-mode original junto con una imagen B-mode con columnas impares interpoladas linealmente.

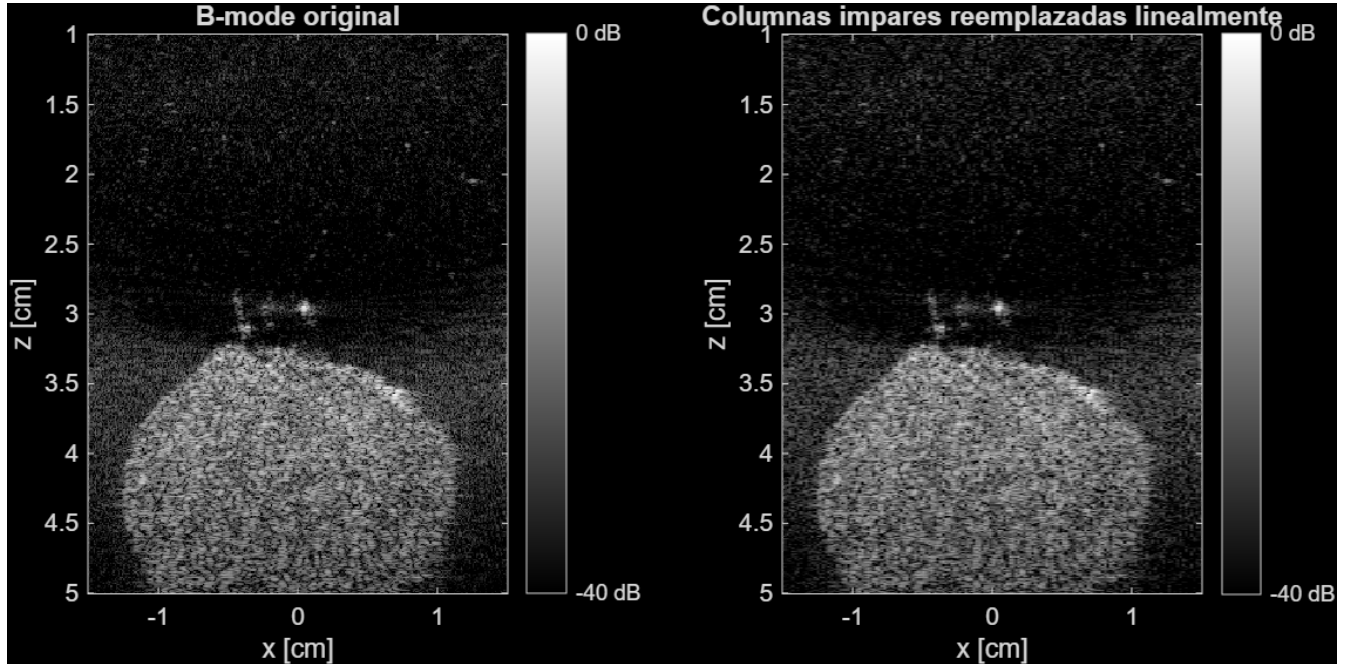


Figura 5.7: Imagen B-mode original y con columnas impares interpoladas linealmente

A continuación se presenta la imagen B-mode generada al reemplazar sus columnas impares por columnas interpoladas mediante la solución analítica de transporte óptimo unidimensional entre columnas vecinas.

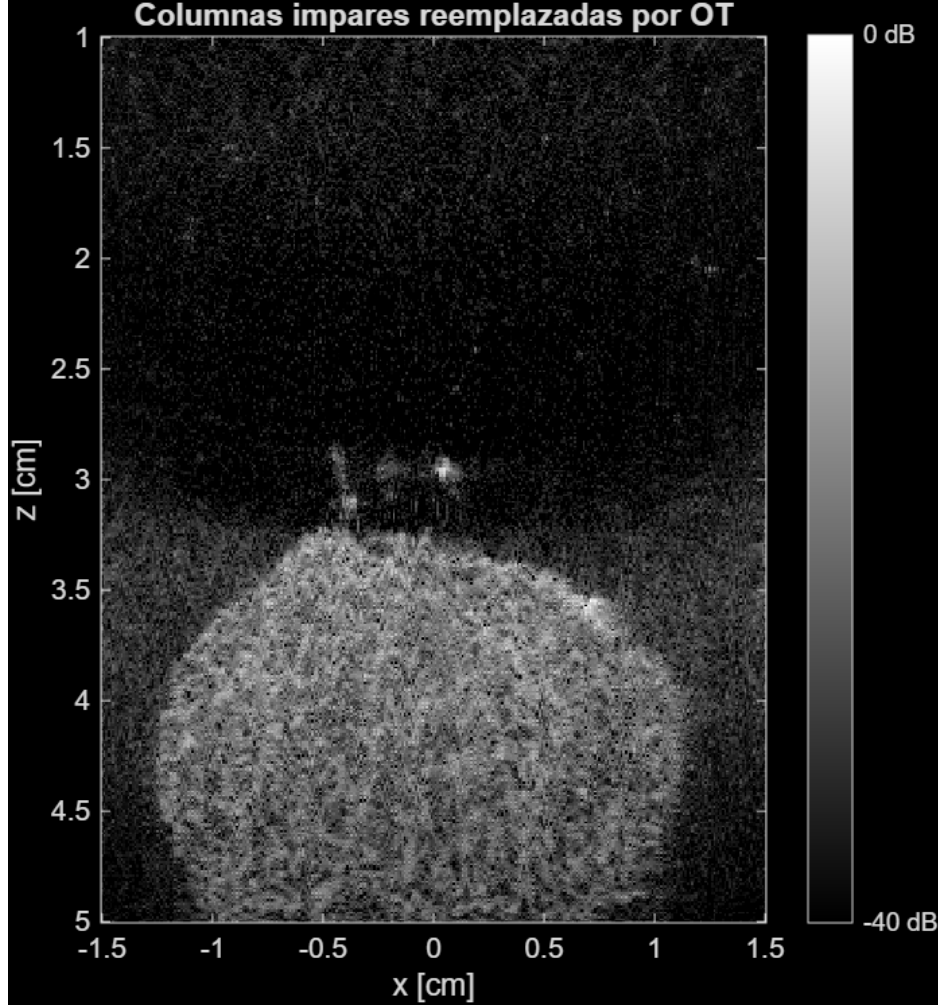


Figura 5.8: Imagen B-mode con columnas impares reemplazadas por OT unidimensional

En las siguientes figuras se muestran las imagenes B-mode generadas al reemplazar columnas impares por interpolaciones mediante UOT entre columnas vecinas. Recordar que el factor ρ de la formulación desbalanceada del problema de transporte óptimo regula la penalización por violar las restricciones en las marginales del plan de transporte π , donde un factor $\rho \rightarrow \infty$ recupera el problema OT balanceado, mientras que $\rho \rightarrow 0$ permite "crear" y "destruir" masa en lugar de transportarla a grandes distancias.

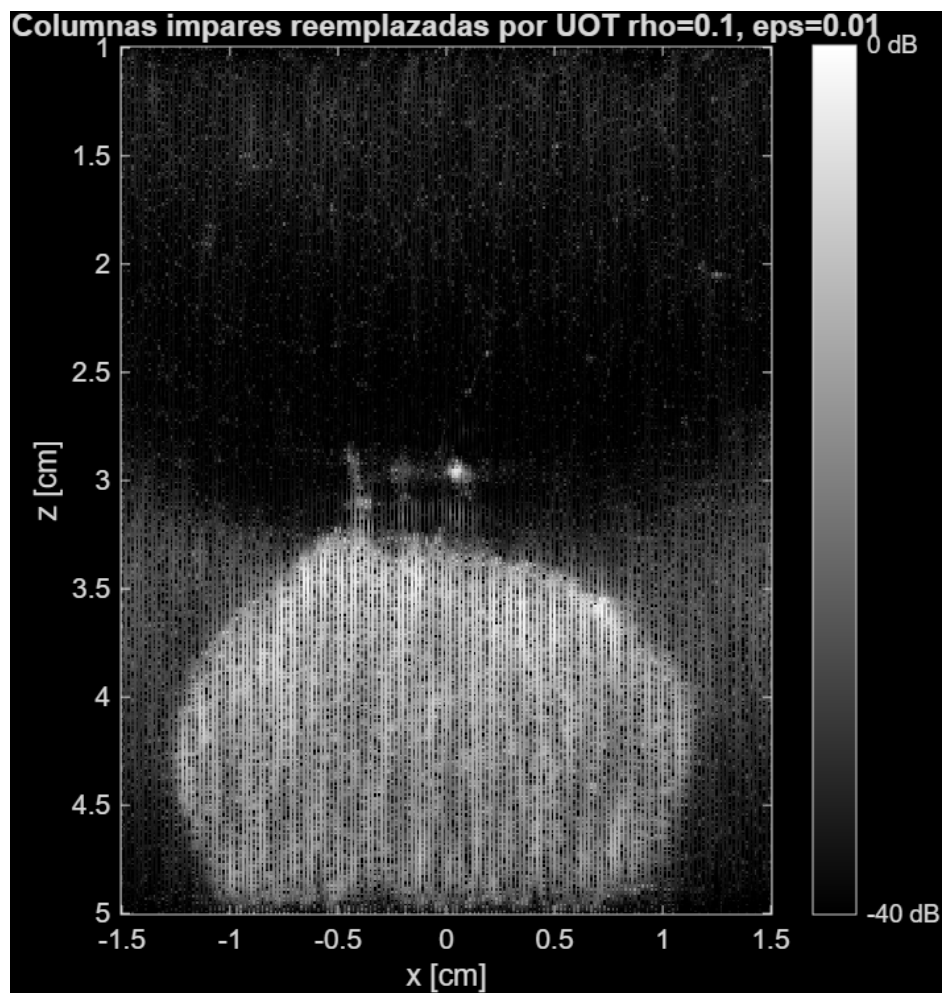


Figura 5.9: Imagen B-mode con columnas impares reemplazadas por UOT $\rho = 0.1$

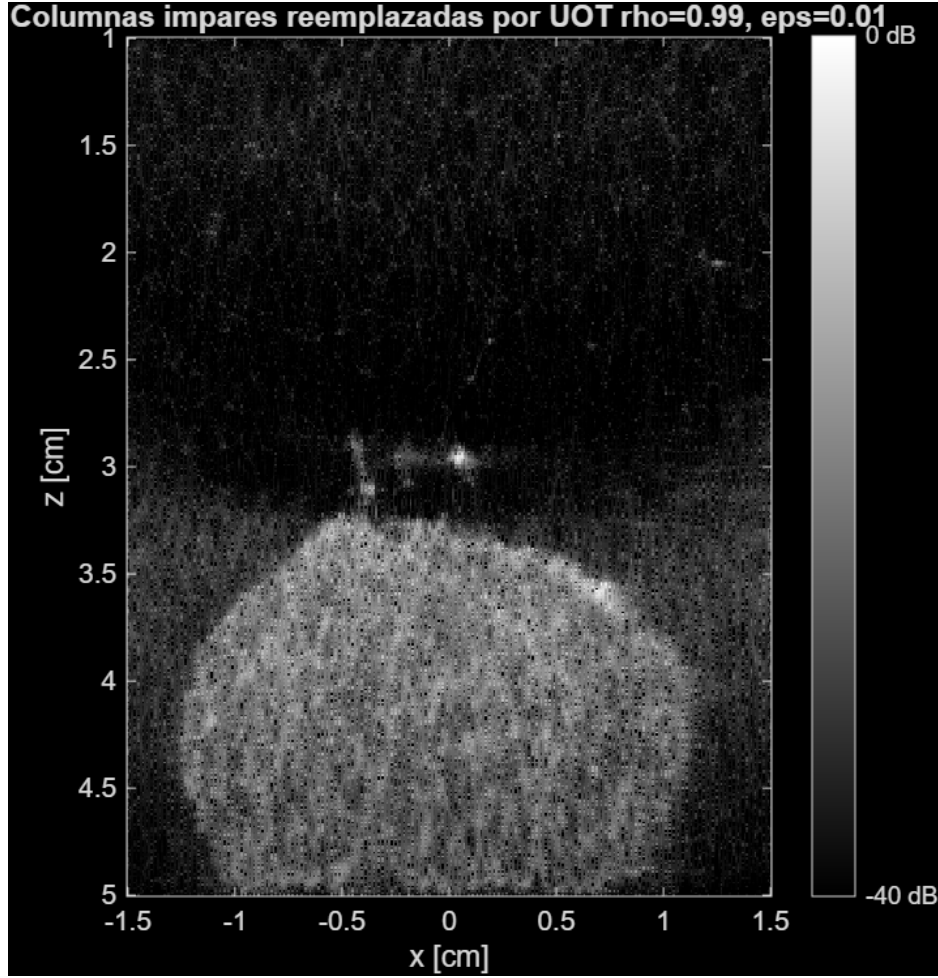


Figura 5.10: Imagen B-mode con columnas impares reemplazadas por UOT $\rho = 0.99$

5.2.1 Análisis de resultados

En las figuras 5.7, 5.8, 5.9, 5.10 se puede apreciar que todos los metodos de interpolación aplicados sobre las columnas de la imagen B-mode original ocasionan una difuminación en el eje x, siendo la interpolación lineal la que genera la difuminación mas leve y la interpolación UOT con factor $\rho = 0.1$ la que genera la difuminación mas exagerada. Estas difuminaciones se ven cuantificadas de forma clara mediante el error RMSE entre la columna interpolada y su respectiva columna en la imagen B-mode original, adicionalmente se grafica tambien las distancias de Wasserstein orden 2.

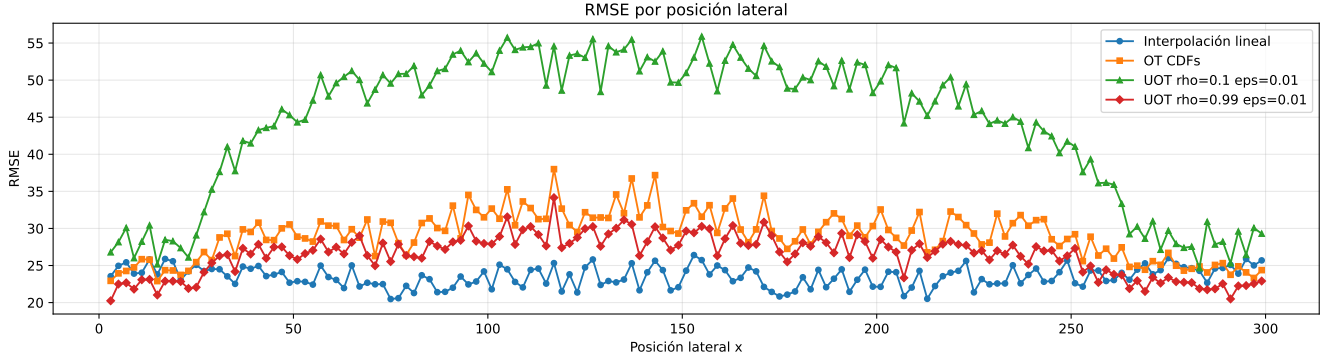


Figura 5.11: RMSE entre columnas interpoladas y originales

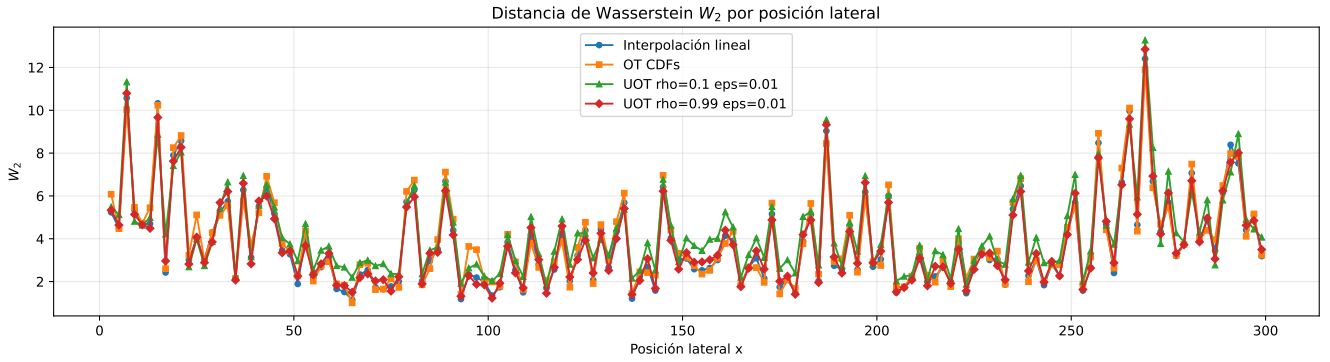


Figura 5.12: Distancia de Wasserstein orden 2 entre columnas interpoladas y originales

En la figura 5.12 podemos ver que todos los metodos tienen una distancia de Wasserstein de orden 2 con respecto a las columnas de la imagen B-mode original similares. Como se puede observar en 5.7, la imagen B-mode original presenta una gran cantidad de pixeles con intensidad igual a 0 en el rango de profundidad entre $z = 3[cm]$ y $z = 1[cm]$, esto sumado al hecho del relativamente bajo numero de muestras que componen cada columna permite que las columnas interpoladas por distintos metodos no sean tan costosos de reordenar para conformar la columna de la imagen B-mode original correspondiente, entregando asi valores similares en la distancias W^2 . Este efecto de reordenamiento se puede ver de forma mas explicita en la figura 5.8 donde se puede ver un leve desplazamiento vertical en los pixeles de cada columna. Si bien algunos de estos efectos de difuminación basados en interpolaciones OT suavizan las imagenes B-mode, a diferencia de las señales RF, no existe un desplazamiento subyacente entre los pixeles que componen cada columna, por lo que las interpolaciones realizadas no son poseen una coherencia fisica dentro de la imagen.

5.3 Interpolación sobre matrices B-mode

Para esta sección se seleccionó una matriz B-mode correspondiente a una toma de datos modalidad quasi-estática para que movimientos muy bruscos no se produzcan diferencias muy exageradas entre los frames interpolados y los originales. Estas matrices B-mode poseen unas dimensiones [529, 542, 97], siendo las primeras dos dimensiones las que definen el tamaño de cada frame y la última dimensión la que define el número de frames totales que componen la matriz. Dada la dificultad de plasmar el movimiento visual que las 97 frames de la matriz producen se adjunta un enlace que muestra un video del dato original sobre el cual se trabaja. (<https://youtu.be/XehxXzCxGi0>)

Como se explico en 4.3, el cálculo del plan de transporte constituye la etapa más costosa computacionalmente. Si bien el uso de una metodología basada en bloques y reconstrucción mediante Overlap-ADD (OLA) permite reducir considerablemente el costo al comparar el necesario de resolver el problema sobre un par de frames completos, cada interpolación continúa requiriendo la resolución de un gran número de problemas de transporte óptimo locales, uno por cada ventana utilizada en la descomposición de un frame.

Con el objetivo de disminuir el número total de problemas por resolver, se optó por calcular un único plan de transporte entre parejas de frames separados por cuatro instantes temporales y utilizarlo para generar tres interpolaciones intermedias. En particular, se calcula el plan de transporte entre los frames 1 y 5, a partir del cual se obtienen los frames interpolados 2, 3 y 4, con valores $\alpha = 0.25$, $\alpha = 0.50$, $\alpha = 0.75$. Posteriormente se realiza el mismo procedimiento entre los frames 5 y 9 para generar los frames 6, 7, 8, y así sucesivamente hasta llegar al frame 97.

Esta estrategia permite reducir en factor de cuatro el número de problemas de transporte óptimo por resolver y además permite analizar de mejor manera el efecto del parámetro α sobre la evolución de las imágenes interpoladas.

Para el método Overlap-Add (OLA) se utilizaron los siguientes parámetros:

- Función ventana: Hamming
- Tamaño parche: 24x24
- Solapamiento entre parches: %50

A continuación se muestran los enlaces que llevan a los videos de las distintas interpolaciones realizadas sobre los datos originales:

- Sinkhorn $\epsilon = 0.1$: <https://youtu.be/0NFoKbsHYbQ>
- Sinkhorn $\epsilon = 1$: https://youtu.be/x040UCfk_RU
- Sinkhorn $\epsilon = 5$: <https://youtu.be/3g7PxVDh0Wc>

- Lineal: https://youtu.be/mJ4lqxM_dTI

5.3.1 Análisis de resultados

Para el análisis de resultados de esta sección, además de estudiar la evolución del error mediante la raíz del error cuadrático medio (RMSE), se emplearán las métricas Peak Signal-to-Noise Ratio (PSNR) y Structural Similarity Index Measure (SSIM), las cuales son explicadas en profundidad en el artículo "Mean Squared Error: Love It or Leave It? A New Look at Signal Fidelity Measures" [9].

La métrica PSNR es una medida de fidelidad derivada del error cuadrático medio (MSE), la cual cuantifica la diferencia entre una imagen de referencia y una imagen reconstruida. Se expresa en decibeles (dB) y se define como

$$PSNR = 10 \log_{10} \left(\frac{L^2}{MSE} \right), \quad (5.1)$$

donde L representa el rango dinámico de intensidades de la imagen. Valores elevados de PSNR indican una mayor similitud entre ambas imágenes [9].

Por otra parte, el índice SSIM evalúa la similitud estructural entre dos imágenes considerando simultáneamente diferencias de luminancia, contraste y estructura local. A diferencia de métricas basadas únicamente en errores punto a punto, SSIM busca aproximar la percepción visual humana de la calidad de una imagen. Su expresión está dada por

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (5.2)$$

donde μ_x y μ_y representan las medias locales de las imágenes, σ_x^2 y σ_y^2 sus varianzas, σ_{xy} la covarianza local, y C_1 , C_2 son constantes utilizadas para estabilizar la división cuando los denominadores son pequeños. El índice SSIM toma valores en el intervalo $[-1, 1]$, aunque en aplicaciones de procesamiento de imágenes normalmente se encuentra entre 0 y 1, siendo este último el valor correspondiente a dos imágenes idénticas [9].

A continuación se presenta el gráfico del error RMSE, PSNR y SSIM de los frames interpolados con respecto a los datos originales.

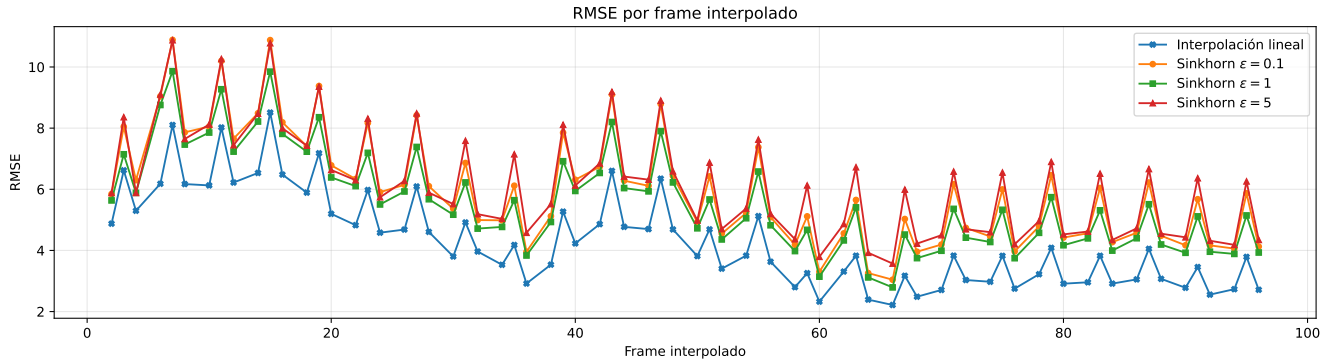


Figura 5.13: RMSE entre frames interpoladas y originales

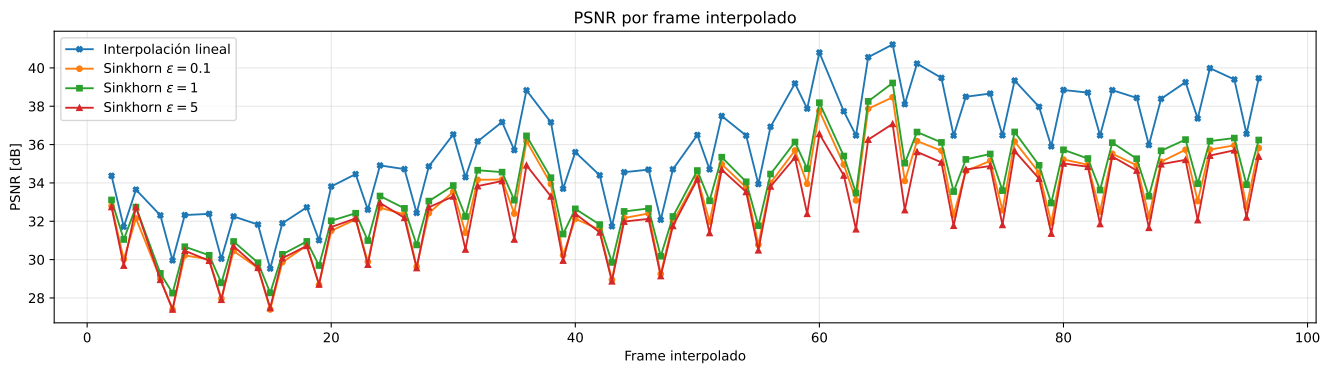


Figura 5.14: PSNR entre frames interpoladas y originales

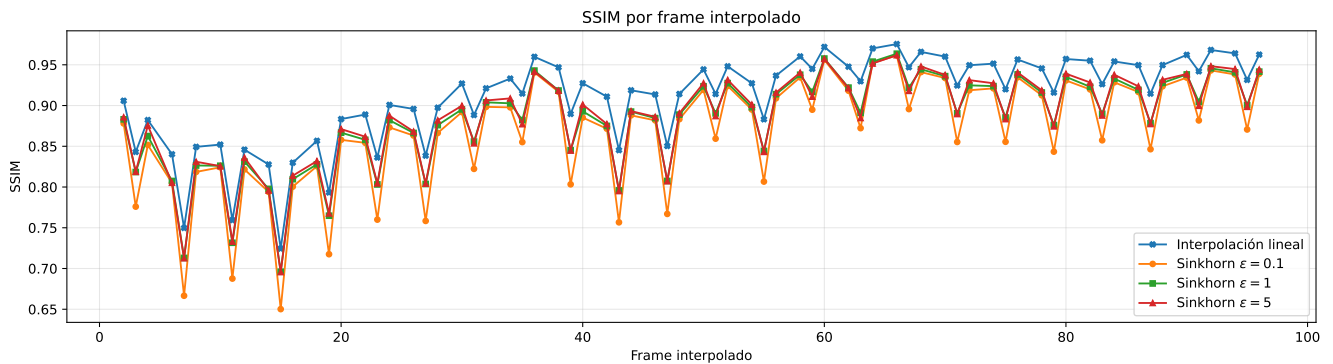


Figura 5.15: SSIM entre frames interpoladas y originales

En los videos entregados puede observarse que la interpolación lineal produce un comportamiento similar al mostrado en la Figura 3.10. En este caso, las intensidades no parecen desplazarse de una posición a otra, sino que disminuyen progresivamente en la ubicación inicial mientras aumentan en la ubicación final. Como consecuencia, los máximos aparentan surgir y desaparecer desde el plano

base de la representación bidimensional, dando la impresión de que la masa se crea y destruye localmente en lugar de trasladarse de manera continua. En contraste, las interpolaciones basadas en transporte óptimo preservan la identidad geométrica de la distribución. La masa no aparece ni desaparece localmente, sino que sigue trayectorias determinadas por el plan de transporte, lo que genera la percepción de que los picos se desplazan como una estructura coherente a lo largo del dominio. A medida que aumenta el parámetro de regularización entrópica ϵ , el plan de transporte se vuelve más difuso, por lo que la masa deja de desplazarse por trayectorias bien definidas y comienza a repartirse entre múltiples destinos, suavizando visualmente la interpolación.

Si bien en los videos el desplazamiento de estructuras geometricas se ve de forma mas clara en las interpolaciones OT, las figuras 5.13, 5.14 y 5.15 muestran nuevamente que la interpolación lineal tiene una pequeña ventaja sobre todas las interpolaciones basadas en OT. Estas pequeñas diferencias que se aprecian son producto del proceso de normalización y posterior reescalamiento que se realiza para poder interpolar mediante transporte óptimo, dado que estos metotods exigen una correspondencia de masa total en ambas distribuciones de origen y destino. De estas figuras tambien se puede apreciar que todos los metodos de interpolación son sensibles a cambios muy bruscos en las imagenes, en el caso de la matriz B-mode original los cambios mas bruscos se producen en los primeros 20 frames, es por eso que todos los graficos transmiten que estos frames interpolados no conforman una representación fidedigna de los datos originales.

Capítulo 6

Conclusiones y Recomendaciones

En este trabajo se exploró la aplicación de técnicas de interpolación basadas en transporte óptimo en datos ultrasónicos pertenecientes a distintas etapas del proceso de construcción de imágenes B-mode, como son las señales RF, envolturas post-beamforming y la propia imagen B-mode. Los resultados obtenidos demuestran que el transporte óptimo permite generar interpolaciones que conservan la información geométrica, plasmando desplazamientos de las estructuras en lugar de producir la aparición y desaparición local de amplitudes. No obstante, aún es necesario ajustar algunos aspectos para que estas ventajas se traduzcan en una mejora de las métricas cuantitativas utilizadas con respecto a la interpolación, como se puede apreciar en todos los gráficos de comparación presentados en el capítulo de Resultados. En particular, el proceso de normalización requerido para formular el problema de transporte óptimo y posterior reescalado de las distribuciones constituyen una importante fuente de error que representa uno de los principales desafíos a abordar en trabajos futuros.

En cuanto a observaciones y recomendaciones acerca de la aplicación de interpolaciones en distintas etapas del proceso de generación de imagen B-mode. En el caso de la interpolación de señales RF, una posible línea de investigación consiste en reemplazar la representación espectral basada en la STFT por la transformada Wavelet, que podría ofrecer una representación tiempo-frecuencial más rigurosa sobre la cual trabajar. Asimismo, resulta interesante el desarrollo de funciones costo que incorporen modelos físicos de propagación ultrasónica, permitiendo que el transporte de masa refleje relaciones físicamente coherentes entre las señales para no perjudicar el proceso de beamforming DAS. Respecto a la interpolación entre columnas de imágenes B-mode, si bien los métodos aplicados producen un efecto visual de suavizado agradable y mejor la continuidad entre columnas adyacentes, su interpretación física es limitada, ya que la interpolación corresponde principalmente a un reordenamiento de intensidades entre columnas y no necesariamente representa la formación real de una adquisición intermedia.

En conjunto, los resultados obtenidos muestran que el transporte óptimo constituye una

herramienta con un potencial considerable para la interpolación de datos ultrasónicos; sin embargo, su aplicación práctica requiere incorporar restricciones y modelos específicos del problema físico para aprovechar plenamente las ventajas geométricas que ofrece esta formulación.

Referencias

- [1] Thomas L. Szabo and Peter Kaczowski. *Essentials of Ultrasound Imaging*. Academic Press, 2024. An imprint of Elsevier.
- [2] François Varray Damien Garcia Vincent Perrot, Maxime Polichetti. So you think you can do? a viewpoint on delay-and-sum beamforming. *Ultrasonics*, 111:2–4, 2021.
- [3] Gaspard Monge. Mémoire sur la théorie des déblais et des remblais. *Histoire de l’Académie Royale des Sciences de Paris*, 1781.
- [4] Leonid V. Kantorovich. On the translocation of masses. *C. R. (Doklady) Acad. Sci. URSS (N.S.)*, 37(7–8):199–201, 1942.
- [5] Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, volume 26, pages 2292–2300, 2013.
- [6] L  na  c Chizat, Gabriel Peyr  , Bernhard Schmitzer, and Fran  ois-Xavier Vialard. Unbalanced optimal transport: Dynamic and kantorovich formulations. *Journal of Functional Analysis*, 274(11):3090–3123, 2018.
- [7] Trevor Henderson and Justin Solomon. Audio transport: A generalized portamento via optimal transport. In *Proceedings of the 22nd International Conference on Digital Audio Effects (DAFx-19)*, pages 1–5, Birmingham, United Kingdom, September 2019.
- [8] Matthew Thorpe, Serim Park, Soheil Kolouri, Gustavo K. Rohde, and Dejan Slep  ev. A transportation l^p distance for signal analysis. *Mathematical Models and Methods in Applied Sciences*, 27(9):1693–1721, 2017.
- [9] Zhou Wang and Alan C. Bovik. Mean squared error: Love it or leave it? a new look at signal fidelity measures. *IEEE Signal Processing Magazine*, 26(1):98–117, 2009.

Apéndice A

Códigos metodología

Listing A.1: Función para encontrar plan de transporte óptimo entre dos distribuciones discretas de probabilidad a y b con matriz de costos C , mediante iteraciones de Sinkhorn

```

1 def sinkhorn_plan(a, b, C, epsilon, n_iter, tol):
2
3     a = np.asarray(a, dtype=np.float64)
4     b = np.asarray(b, dtype=np.float64)
5     C = np.asarray(C, dtype=np.float64)
6
7     a = a / a.sum()
8     b = b / b.sum()
9
10    K = np.exp(-C / epsilon)
11
12    u = np.ones_like(a)
13    v = np.ones_like(b)
14
15    for it in range(n_iter):
16
17        u = a / (K @ v + 1e-16)
18        v = b / (K.T @ u + 1e-16)
19
20        P = (u[:, None] * K) * v[None, :]
21
22        err_a = np.linalg.norm(P.sum(axis=1) - a, ord=1)
23        err_b = np.linalg.norm(P.sum(axis=0) - b, ord=1)
24
25        if max(err_a, err_b) < tol:
26            break
27
28    return P

```

Listing A.2: Función para interpolar una distribución baricéntrica según un plan de transporte

```

1 def interpolar_desde_plan(P, k=0.5):
2
3     P = np.asarray(P, dtype=np.float64)
4     n, m = P.shape
5
6     interpolada = np.zeros(n)

```

```

7
8     for i in range(n):
9         for j in range(n):
10             pos = int(round((1 - k) * i + k * j))
11             interpolada[pos] += P[i, j]
12
13     return interpolada

```

Listing A.3: Construcción del espectro utilizado para transporte óptimo mediante STFT

```

1 def espectro_OT(senal, fs, window, nperseg, noverlap):
2
3     f, t, Zxx = stft(senal, window, fs, nperseg, noverlap)
4
5     mag = np.abs(Zxx)
6
7     mag_colnorm = Zxx/np.sum(Zxx, axis=0)
8     columnas_normalizadas = np.array(mag_colnorm)
9
10    normas_columnas = np.sum(mag, axis=0)
11
12    return Zxx, normas_columnas, columnas_normalizadas

```

Listing A.4: Reconstrucción de fase mediante acumulación de frecuencia instantánea tipo phase vocoder

```

1
2 def wrap_phase(x):
3     return np.angle(np.exp(1j * x))
4
5
6 def frecuencia_instantanea(Zxx, fs, nperseg, noverlap):
7
8     hop = nperseg - noverlap
9     dt = hop / fs
10
11    n_freqs, n_frames = Zxx.shape
12
13    f = np.fft.rfftfreq(nperseg, d=1/fs)
14    omega_bins = 2 * np.pi * f

```

```

15
16     fase = np.angle(Zxx)
17
18     omega_inst = np.zeros_like(fase)
19     omega_inst[:, 0] = omega_bins
20
21     for t in range(1, n_frames):
22
23         dphi = wrap_phase(fase[:, t] - fase[:, t-1])
24
25         esperado = omega_bins * dt
26
27         desviacion = wrap_phase(dphi - esperado)
28
29         omega_inst[:, t] = omega_bins + desviacion / dt
30
31     return omega_inst
32
33
34 def reconstruir_fase_phase_vocoder(Zxx_1, Zxx_2, k, fs, nperseg, noverlap):
35
36     hop = nperseg - noverlap
37     dt = hop / fs
38
39     omega_1 = frecuencia_instantanea(Zxx_1, fs, nperseg, noverlap)
40
41     omega_2 = frecuencia_instantanea(Zxx_2, fs, nperseg, noverlap)
42
43     omega_k = (1 - k) * omega_1 + k * omega_2
44
45     fase_1 = np.angle(Zxx_1[:, 0])
46     fase_2 = np.angle(Zxx_2[:, 0])
47
48     fase_k = np.zeros_like(np.angle(Zxx_1))
49
50     fase_k[:, 0] = np.angle(
51         (1 - k) * np.exp(1j * fase_1)
52         + k * np.exp(1j * fase_2)

```

```

53     )
54
55     for t in range(1, Zxx_1.shape[1]):
56
57         fase_k[:, t] = (
58             fase_k[:, t-1]
59             + omega_k[:, t-1] * dt
60         )
61
62     return fase_k

```

Listing A.5: Interpolación espectral mediante transporte óptimo y reconstrucción de fase tipo phase vocoder

```

1  def interpolarseales(se al_1, se al_2, C, k):
2
3      Zxx_1, norms_1, cols_1 = espectro_OT(se al_1)
4      Zxx_2, norms_2, cols_2 = espectro_OT(se al_2)
5
6      n = Zxx_1.shape[0]
7
8      magnitudes_interpoladas = np.zeros_like(cols_1)
9
10     for t in range(Zxx_1.shape[1]):
11
12         a = cols_1[:, t]
13         b = cols_2[:, t]
14
15         plan = sinkhorn_plan(a, b, C, epsilon, n_iter, tol)
16         interpolacion_columnas = interpolardesdeplan(plan, k)
17
18         norma_interpolada = (
19             (1 - k) * norms_1[t]
20             + k * norms_2[t]
21         )
22
23         interpolacion_columnas *= norma_interpolada
24
25     magnitudes_interpoladas = (

```

```

26         magnitudes_interpoladas
27         .at[:, t]
28         .set(interpolacion_columnas)
29     )
30
31     fase_interpolada = reconstruir_fase_phase_vocoder(
32         np.asarray(Zxx_1),
33         np.asarray(Zxx_2),
34         k,
35         fs,
36         nperseg,
37         noverlap=None
38     )
39
40     Zxx_interpolado = (
41         np.asarray(magnitudes_interpoladas)
42         * np.exp(1j * fase_interpolada)
43     )
44
45     _, se_al_interpolada = istft(
46         Zxx_interpolado,
47         fs,
48         nperseg,
49         noverlap
50     )
51
52     return se_al_interpolada

```

Listing A.6: Interpolación mediante distancia TL^p

```

1 def tlp_interpolate(f, g, alpha, lam, epsilon, p, n_iter):
2
3     f = np.asarray(f, dtype=np.float64).ravel()
4     g = np.asarray(g, dtype=np.float64).ravel()
5
6     N = len(f)
7     t = np.linspace(0, 1, N)
8
9     a = np.ones(N) / N

```

```

10     b = np.ones(N) / N
11
12     scale = max(np.max(np.abs(f)), np.max(np.abs(g)), 1e-12)
13     f0 = f / scale
14     g0 = g / scale
15
16     Ct = np.abs(t[:, None] - t[None, :]) ** p
17     Ca = np.abs(f0[:, None] - g0[None, :]) ** p
18     C = (Ct / lam) + Ca
19
20     P = sinkhorn(a, b, C, epsilon=epsilon, n_iter=n_iter)
21
22     t_mid = (1 - alpha) * t[:, None] + alpha * t[None, :]
23     s_mid = (1 - alpha) * f0[:, None] + alpha * g0[None, :]
24
25     out = np.zeros(N)
26     w = np.zeros(N)
27
28     idx = np.round(t_mid * (N - 1)).astype(int)
29     idx = np.clip(idx, 0, N - 1)
30
31     np.add.at(out, idx.ravel(), (P * s_mid).ravel())
32     np.add.at(w, idx.ravel(), P.ravel())
33
34     out = out / (w + 1e-12)
35     out = out * scale
36
37     return out, P, C

```

Listing A.7: Interpolación de señales RF mediante bloques y overlap-add

```

1 def tlp_interpolate_signal_blocks(signal1, signal2, alpha,
2                                   block_length, hop_length, window,
3                                   lam, epsilon, p, n_iter):
4
5     signal1 = np.asarray(signal1, dtype=np.float64).ravel()
6     signal2 = np.asarray(signal2, dtype=np.float64).ravel()
7
8     N = len(signal1)

```

```

9
10 if N <= block_length:
11     n_blocks = 1
12 else:
13     n_blocks = int(np.ceil((N - block_length) / hop_length)) + 1
14
15 total_len = ((n_blocks - 1) * hop_length + block_length)
16
17 pad = max(0, total_len - N)
18
19 x1 = np.pad(signal1, (0, pad), mode="constant")
20 x2 = np.pad(signal2, (0, pad), mode="constant")
21
22 win = get_window(window, block_length, fftbins=False).astype(np.float64)
23
24 ola_num = np.zeros(total_len)
25 ola_den = np.zeros(total_len)
26
27 for m in range(n_blocks):
28
29     start = m * hop_length
30     end = start + block_length
31
32     block1 = x1[start:end]
33     block2 = x2[start:end]
34
35     block_interp, P, C = tlp_interpolate(
36         block1,
37         block2,
38         alpha=alpha,
39         lam=lam,
40         epsilon=epsilon,
41         p=p,
42         n_iter=n_iter,
43     )
44
45     ola_num[start:end] += block_interp * win
46     ola_den[start:end] += win

```

```

47
48     signal_interp = ola_num / (ola_den + 1e-12)
49     signal_interp = signal_interp[:N]
50
51     return signal_interp

```

Listing A.8: Interpolación de columnas mediante UOT y baricentro del plan de transporte

```

1  function col_mid = uot_interp_col(col1, col2, alpha, opts)
2
3  if nargin < 4, opts = struct; end
4  opts = setdef(opts, 'epsilon', 1e-3);
5  opts = setdef(opts, 'rho', 0.05);
6  opts = setdef(opts, 'nIter', 500);
7  opts = setdef(opts, 'tol', 1e-7);
8
9
10 a = double(col1(:));
11 b = double(col2(:));
12
13
14 N = numel(a);
15
16 scale = max([sum(a), sum(b), eps]);
17 a0 = a / scale;
18 b0 = b / scale;
19
20 x = linspace(0,1,N)';
21 C = (x - x').^2;
22
23 K = exp(-C / opts.epsilon);
24 K = K + realmin;
25
26 tau = opts.rho / (opts.rho + opts.epsilon);
27
28 u = ones(N,1);
29 v = ones(N,1);
30
31 for it = 1:opts.nIter

```

```

32
33     u_old = u;
34
35     Kv = K * v + realmin;
36     u = (a0 ./ Kv).^tau;
37
38     Ktu = K' * u + realmin;
39     v = (b0 ./ Ktu).^tau;
40
41     if norm(u - u_old) / max(norm(u_old),eps) < opts.tol
42         break
43     end
44
45 end
46
47 P = (u .* K) .* v';
48
49 col_mid0 = zeros(N,1);
50
51 for i = 1:N
52     for j = 1:N
53
54         m = P(i,j);
55
56         if m <= 0
57             continue;
58         end
59
60         pos = (1-alpha)*i + alpha*j;
61
62         k0 = floor(pos);
63         k1 = ceil(pos);
64
65         if k0 < 1
66             col_mid0(1) = col_mid0(1) + m;
67
68         elseif k1 > N
69             col_mid0(N) = col_mid0(N) + m;

```

```

70
71     elseif k0 == k1
72         col_mid0(k0) = col_mid0(k0) + m;
73
74     else
75
76         w1 = pos - k0;
77         w0 = 1 - w1;
78
79         col_mid0(k0) = col_mid0(k0) + w0*m;
80         col_mid0(k1) = col_mid0(k1) + w1*m;
81
82     end
83 end
84 end
85
86 col_mid = col_mid0 * scale;
87
88 end

```

Listing A.9: Interpolación OT 1D mediante CDFs y funciones cuantiles

```

1 function col_mid = ot_cdf_interp_col(col1, col2, alpha)
2
3 a = double(col1(:));
4 b = double(col2(:));
5
6 sa = sum(a);
7 sb = sum(b);
8
9 a = a / sa;
10 b = b / sb;
11
12 N = length(a);
13
14 Fa = cumsum(a);
15 Fb = cumsum(b);
16
17 u = linspace(0,1,N)';

```

```

18
19 Fa_inv = interp1(Fa, (1:N)', u, 'linear', 'extrap');
20 Fb_inv = interp1(Fb, (1:N)', u, 'linear', 'extrap');
21
22 Fmid_inv = (1-alpha)*Fa_inv + alpha*Fb_inv;
23
24
25 col_mid = zeros(N,1);
26
27 for k = 1:N
28
29     pos = Fmid_inv(k);
30
31     i0 = floor(pos);
32     i1 = ceil(pos);
33
34     if i0 < 1
35
36         col_mid(1) = col_mid(1) + 1/N;
37
38     elseif i1 > N
39
40         col_mid(N) = col_mid(N) + 1/N;
41
42     elseif i0 == i1
43
44         col_mid(i0) = col_mid(i0) + 1/N;
45
46     else
47
48         w1 = pos - i0;
49         w0 = 1 - w1;
50
51         col_mid(i0) = col_mid(i0) + w0*(1/N);
52         col_mid(i1) = col_mid(i1) + w1*(1/N);
53
54     end
55

```

```

56 end
57
58 scale = (sa + sb)/2;
59 col_mid = col_mid * scale;
60
61 end

```

Listing A.10: Upsampling lateral sobre envolventes mediante interpolaciones UOT

```

1 function Env_hi = uot_upsample_x(Env, opts, alphas)
2
3     [Nz, Nx] = size(Env);
4     nAlpha = numel(alphas);
5
6     Env_hi = zeros(Nz, (Nx-1)*(nAlpha+1) + 1);
7
8     idx = 1;
9
10    for i = 1:Nx-1
11
12        Env_hi(:,idx) = Env(:,i);
13        idx = idx + 1;
14
15        for aidx = 1:nAlpha
16
17            alpha = alphas(aidx);
18
19            Env_hi(:,idx) = uot_interp_col(Env(:,i), Env(:,i+1), ...
20                                           alpha, opts);
21
22            idx = idx + 1;
23        end
24    end
25
26    Env_hi(:,idx) = Env(:,end);
27
28 end

```

Listing A.11: Upsampling lateral sobre envolventes mediante interpolaciones lineales

```

1 function Env_hi = linear_upsample_x(Env, alphas)
2
3     [Nz, Nx] = size(Env);
4     nAlpha = numel(alphas);
5
6     Env_hi = zeros(Nz, (Nx-1)*(nAlpha+1) + 1);
7
8     idx = 1;
9
10    for i = 1:Nx-1
11
12        Env_hi(:,idx) = Env(:,i);
13        idx = idx + 1;
14
15        for aidx = 1:nAlpha
16
17            alpha = alphas(aidx);
18
19            Env_hi(:,idx) = (1-alpha)*Env(:,i) + alpha*Env(:,i+1);
20
21            idx = idx + 1;
22
23        end
24    end
25
26    Env_hi(:,idx) = Env(:,end);
27
28 end

```

Listing A.12: Interpolación bidimensional mediante Sinkhorn regularizado

```

1 function Barys = sinkhorn2d_patch_multialpha(P1, P2, alphas, epsilon, nIter)
2
3 P1 = double(max(P1,0));
4 P2 = double(max(P2,0));
5
6 [Nz, Nx] = size(P1);
7 n = Nz * Nx;
8 nA = numel(alphas);

```

```

9
10 M1 = sum(P1(:));
11 M2 = sum(P2(:));
12
13 a = P1(:) + 1e-12;
14 b = P2(:) + 1e-12;
15
16 a = a / sum(a);
17 b = b / sum(b);
18
19 [Z, X] = ndgrid(1:Nz, 1:Nx);
20 coords = [Z(:), X(:)];
21
22 C = local_sqdist(coords);
23 K = exp(-C / epsilon);
24
25 u = ones(n,1) / n;
26 v = ones(n,1) / n;
27
28 for it = 1:nIter
29     u = a ./ (K*v + 1e-300);
30     v = b ./ (K'*u + 1e-300);
31 end
32
33 P = (u .* K) .* v';
34
35 zi0 = coords(:,1);
36 xi0 = coords(:,2);
37
38 zi1 = coords(:,1)';
39 xi1 = coords(:,2)';
40
41 Barys = zeros(Nz, Nx, nA);
42
43 for ia = 1:nA
44
45     alpha = alphas(ia);
46

```

```

47     Malpha = (1-alpha)*M1 + alpha*M2;
48     mass = P(:) * Malpha;
49
50     zi = (1-alpha)*zi0 + alpha*zi1;
51     xi = (1-alpha)*xi0 + alpha*xi1;
52
53     zi = zi(:);
54     xi = xi(:);
55
56     z0 = floor(zi);
57     z1 = z0 + 1;
58
59     x0 = floor(xi);
60     x1 = x0 + 1;
61
62     wz1 = zi - z0;
63     wz0 = 1 - wz1;
64
65     wx1 = xi - x0;
66     wx0 = 1 - wx1;
67
68     tmp = zeros(n,1);
69
70     lin00 = sub2ind([Nz Nx], z0, x0);
71     lin01 = sub2ind([Nz Nx], z0, x1);
72     lin10 = sub2ind([Nz Nx], z1, x0);
73     lin11 = sub2ind([Nz Nx], z1, x1);
74
75     tmp = tmp + accumarray(lin00, mass .* wz0 .* wx0, [n 1], @sum, 0);
76     tmp = tmp + accumarray(lin01, mass .* wz0 .* wx1, [n 1], @sum, 0);
77     tmp = tmp + accumarray(lin10, mass .* wz1 .* wx0, [n 1], @sum, 0);
78     tmp = tmp + accumarray(lin11, mass .* wz1 .* wx1, [n 1], @sum, 0);
79
80     Barys(:, :, ia) = reshape(tmp, Nz, Nx);
81
82 end
83
84 end

```

Listing A.13: Interpolación de frames B-mode mediante Sinkhorn 2D por bloques y reconstrucción OLA

```

1 function Finterps = sinkhorn2d_frame_blocks_ola_multialpha(F1, F2, alphas, opts)
2
3 if nargin < 4
4     opts = struct;
5 end
6
7 opts = setdef(opts, 'patch_size', 24);
8 opts = setdef(opts, 'hop', 12);
9 opts = setdef(opts, 'epsilon', 5);
10 opts = setdef(opts, 'nIter', 100);
11
12 F1 = double(F1);
13 F2 = double(F2);
14
15 [Nz, Nx] = size(F1);
16
17 ps = opts.patch_size;
18 hop = opts.hop;
19 nA = numel(alphas);
20
21 w = hann(ps, 'periodic');
22 W = w * w';
23
24 starts_z = 1:hop:(Nz-ps+1);
25 starts_x = 1:hop:(Nx-ps+1);
26
27 if starts_z(end) ~= Nz-ps+1
28     starts_z = [starts_z, Nz-ps+1];
29 end
30
31 if starts_x(end) ~= Nx-ps+1
32     starts_x = [starts_x, Nx-ps+1];
33 end
34
35 ola_num = zeros(Nz, Nx, nA);
36 ola_den = zeros(Nz, Nx);

```

```

37
38 for iz = 1:numel(starts_z)
39
40     z0 = starts_z(iz);
41
42     for ix = 1:numel(starts_x)
43
44         x0 = starts_x(ix);
45
46         P1 = max(F1(z0:z0+ps-1, x0:x0+ps-1), 0) .* W;
47         P2 = max(F2(z0:z0+ps-1, x0:x0+ps-1), 0) .* W;
48
49         Barys = sinkhorn2d_patch_multialpha( ...
50             P1, P2, alphas, opts.epsilon, opts.nIter);
51
52         for ia = 1:nA
53             ola_num(z0:z0+ps-1, x0:x0+ps-1, ia) = ...
54                 ola_num(z0:z0+ps-1, x0:x0+ps-1, ia) + Barys(:,:,ia) .* W;
55         end
56
57         ola_den(z0:z0+ps-1, x0:x0+ps-1) = ...
58             ola_den(z0:z0+ps-1, x0:x0+ps-1) + W.^2;
59
60     end
61 end
62
63 Finterps = zeros(Nz, Nx, nA);
64
65 for ia = 1:nA
66     Finterps(:,:,ia) = ola_num(:,:,ia) ./ (ola_den + 1e-12);
67     Finterps(:,:,ia) = max(Finterps(:,:,ia), 0);
68 end
69
70 end

```