



Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

Generación de vórtices ópticos usando el algoritmo de
Gerchberg-Saxton

Tesis presentada al

Colegio de Física

como requisito parcial para la obtención del grado de

LICENCIADO EN FÍSICA APLICADA

por

Montserrat Huerta Sandoval

Asesorada por

Dr. Miguel Ángel Olvera Santamaría

Dr. Andrey S. Ostrovsky

Puebla Pue.
19 de agosto de 2022



Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

Generación de vórtices ópticos usando el algoritmo de
Gerchberg-Saxton

Tesis presentada al

Colegio de Física

como requisito parcial para la obtención del grado de

LICENCIADO EN FÍSICA APLICADA

por

Montserrat Huerta Sandoval

Asesorada por

Dr. Miguel Ángel Olvera Santamaría

Dr. Andrey S. Ostrovsky

Puebla Pue.
19 de agosto de 2022

Título: Generación de vórtices ópticos usando el algoritmo de Gerchberg-Saxton

Estudiante: MONTSERRAT HUERTA SANDOVAL

COMITÉ

Dr. Carlos I. Robledo Sánchez
Presidente

Dr. Alberto Cordero Dávila
Secretario

Dra. Rosario Pastrana Sánchez
Vocal

Dr. Cruz Meneses Fabian
Vocal

Dr. Miguel Ángel Olvera Santamaría
Dr. Andrey S. Ostrovsky
Asesor

Índice general

Resumen	1
Introducción	3
1. El problema de fase en Óptica	5
1.1. Fundamentos de la teoría escalar de difracción	5
1.1.1. Propagación de un frente de onda	6
1.1.2. Principio de Huygens-Fresnel	8
1.1.3. Formulación de difracción de Rayleigh-Sommerfeld	9
1.1.4. Aproximación de Fresnel	9
1.2. Medición de un campo óptico como una proyección	11
1.2.1. Detección de amplitud y fase	11
1.2.2. Recuperación de fase y factibilidad	11
1.3. Proyecciones en espacios convexos y no convexos	12
1.3.1. Distancia, proyecciones y proyectores	12
1.3.2. Convexidad y cerradura	13
2. Métodos iterativos de Fourier basados en el algoritmo de Gerchberg-Saxton	15
2.1. Transformada y transformada rápida de Fourier	15
2.1.1. Propiedades	15
2.1.2. Discretización de la Transformada y Transformada Inversa de Fourier	16
2.1.3. Transformada rápida de Fourier en MATLAB (FFT)	17
2.2. Algoritmo de Gerchberg-Saxton	17
2.3. Estructura de algoritmos iterativos para recuperación de fase	18
3. Generación de un vórtice óptico usando técnicas de recuperación de fase	21
3.1. Vórtice óptico: definición y características principales	21
3.1.1. Momento angular orbital	21
3.1.2. Vórtice Óptico	21
3.2. Esquemas de generación de vórtices ópticos: coherentes y parcialmente coherentes	23
3.2.1. Vórtices parcialmente coherentes	26
3.3. Síntesis iterativa de un vórtice óptico	29
3.3.1. Análisis de error	30
4. Resultados numéricos	33
4.1. Vórtice óptico perfecto	33
4.2. Vórtice óptico parcialmente coherente	42
4.3. Error cuadrático medio	43
5. Conclusiones	49

Resumen

En este trabajo se propone una técnica para generar vórtices ópticos tanto completamente coherentes como parcialmente coherentes. Dicha técnica se basa en la implementación numérica del algoritmo de Gerchberg-Saxton, el cuál nos permite obtener la fase que genera los vórtices ópticos mencionados a partir de mediciones de sus intensidades. El algoritmo obtenido es analizado para distintos casos en los que puede tenerse una aproximación de la fase buscada.

Introducción

Un vórtice óptico es un haz cuyo frente de onda se propaga en forma helicoidal a lo largo de la dirección de propagación del haz[1]. Una característica fundamental de los vórtices ópticos es su perfil espacial en forma de anillos concéntricos con una singularidad de fase caracterizada por la expresión $\exp(i\nu\phi)$, en donde ϕ es el ángulo azimutal y ν es un número entero llamado carga topológica. Como consecuencia del flujo de energía electromagnética en forma helicoidal, los vórtices ópticos poseen momento angular orbital, el cual es proporcional a la carga topológica [2]. Ésta propiedad hace que los vórtices ópticos sean especialmente interesantes debido a sus diversas aplicaciones relacionadas al atrapamiento óptico y manipulación de micropartículas.

Los vórtices ópticos pueden generarse en sistemas de transformada de Fourier que incorporan hologramas los cuales contienen información necesaria para sintetizar el vórtice. Las técnicas comunes incluyen el uso de platos de fase [3], axicones [4], moduladores espaciales de luz de cristal líquido (LC-SLM)[5] o incluso la implementación física de ensambles estadísticos como es el caso de los vórtices parcialmente coherentes [6]. Sin embargo, las técnicas comunes requieren de codificaciones basadas en la fase real de la señal de interés, un hecho que puede ser complicado en ocasiones. Otra complicación al sintetizar un campo óptico, por ejemplo un vórtice, proviene del hecho de que al realizar una medición de un campo óptico por medio de algún detector de luz se obtiene únicamente información de la intensidad del campo mientras que la información de fase se pierde durante el proceso de medición. Esto significa que la fase no es medible directamente y no puede emplearse como un dato experimental para sintetizar otro vórtice. Este problema es conocido como el problema de fase y su solución tiene un gran número de aplicaciones en áreas como cristalografía de rayos X, astronomía, diseño y pruebas ópticas, microscopía, entre otros [7,8].

Por otra parte, los algoritmos iterativos tipo Gerchberg-Saxton así como sus variantes [9,10], resultan ser una herramienta útil a la hora de realizar aproximaciones a la fase real del campo óptico de interés a partir de mediciones de intensidad además de alguna otra información conocida del campo. El funcionamiento básico de estos algoritmos se basa esencialmente en la aplicación sistemática de la Transformada de Fourier y la Transformada de Fourier Inversa entre los planos objeto y de Fourier del sistema óptico en cuestión, permitiendo con cada iteración una mejor aproximación a la fase que permite reconstruir el campo óptico.

En este trabajo, implementamos el algoritmo de Gerchber-Saxton en Matlab para obtener una aproximación a la fase que genera un vórtice óptico perfecto y un vórtice óptico parcialmente coherente. El primer caso corresponde a un anillo de luz teóricamente infinitamente delgado. El segundo es el caso opuesto, un anillo de luz espacialmente extendido que resulta de un promedio sobre un ensamble estadístico. Esto permite identificar adecuadamente las constricciones físicas reales en el dominio del objeto y de Fourier para definir numéricamente las constricciones numéricas en el algoritmo. En ambos casos encontramos que el algoritmo converge en menos de 100 iteraciones. Analizamos algunas características de los resultados así como fuentes de error y consecuencias de las constricciones empleadas. Consideramos que los resultados obtenidos marcan nuevas rutas de investigación para técnicas holográficas para generación de vórtices, sistemas de pinzas ópticas y micromanipulación.

Capítulo 1

El problema de fase en Óptica

1.1. Fundamentos de la teoría escalar de difracción

Introducción histórica

La difracción es un fenómeno que juega un papel importante en diversas áreas de la física en las que se involucra la propagación de ondas. Ocurre cuando un frente de onda es obstaculizado por un objeto o bien por alguna abertura que se interpone en su camino. Las primeras observaciones de este fenómeno fueron hechas por Grimaldi y publicadas años después de su fallecimiento, en 1665. Grimaldi demostró que al dejar pasar los rayos de sol en un cuarto oscuro a través de dos trozos de cartón con un orificio cada uno, se percató de que la luz proyectada era una mancha mayor a la esperada si la propagación de la luz fuera rectilínea. Años más tarde, siguiendo las ideas y observaciones de Grimaldi, Huygens expresó la idea intuitiva de que si cada punto en el frente de onda de una perturbación se consideraba una nueva fuente de una perturbación esférica "secundaria", entonces el frente de onda en un instante posterior podría encontrarse construyendo la envolvente de las ondas secundarias, como se ilustra en la Fig. 1.1.

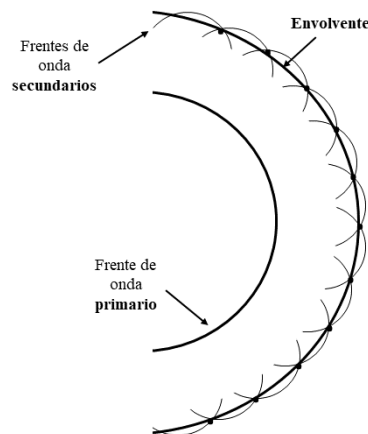


Figura 1.1: Construcción de Huygens: cada punto en el frente de onda de una perturbación puede considerarse como una fuente de ondas esféricas secundarias cuya envolvente constituye el nuevo frente de onda.

Por su parte, Thomas Young introducía el concepto de interferencia en 1803, una idea revolucionaria para su época y que, junto con las ideas de Huygens servirían como impulso para la formulación de Augustin Jean Fresnel, en donde se engloban ambas ideas. Fresnel hizo suposiciones arbitrarias sobre las amplitudes y fases de las fuentes secundarias que proponía Huygens además de la idea de que dichas ondas pueden interferir mutuamente. Está formulación es conocida como *principio de Huygens-Fresnel*.

En 1882, Gustav Kirchhoff logró dar un sustento matemático a las ideas propuestas por Huygens y Fresnel y logró demostrar que las amplitudes y fases atribuidas a las fuentes secundarias por Fresnel eran de hecho consecuencias lógicas de la naturaleza ondulatoria de la luz, un hecho que había sido ya demostrado por Maxwell en 1860. Kirchhoff basó su formulación matemática en dos supuestos sobre los valores límite de la luz incidente en la superficie de un obstáculo colocado en el camino de la propagación de la luz. Sin embargo, a pesar de que hoy en día dicha teoría es bien aceptada, presentaba errores en las hipótesis utilizadas para su deducción que motivaron la búsqueda de un desarrollo matemático más riguroso, un hecho que llegaría años más tarde de la mano de Rayleigh y Sommerfeld con su llamada *teoría de difracción de Rayleigh-Sommerfeld*.

Tanto en la teoría de Kirchhoff como en la teoría de Rayleigh-Sommerfeld, la luz es tratado como un fenómeno escalar, dejando de lado la naturaleza fundamentalmente vectorial de los campos electromagnéticos. Este enfoque tiene en cuenta el hecho de que, en los límites, los diversos componentes de los campos eléctricos y magnéticos están acoplados a través de las ecuaciones de Maxwell y no pueden ser tratados de manera independiente.

1.1.1. Propagación de un frente de onda

Cuando un frente de onda se propaga en el espacio, tanto su fase como su amplitud cambian. Para poder explicar dicha propagación se han desarrollado diversas teorías y modelos que permiten, bajo ciertas condiciones, una descripción de este fenómeno. Partiendo de las ecuaciones de Maxwell en su forma diferencial, en unidades del sistema internacional SI para un medio óptico [11]

$$\nabla \cdot \mathbf{E} = 0, \quad (1.1)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (1.2)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (1.3)$$

$$\nabla \times \mathbf{B} = \mu\epsilon \frac{\partial \mathbf{E}}{\partial t}, \quad (1.4)$$

en donde \mathbf{B} y \mathbf{E} son el campo magnético y eléctrico respectivamente, ϵ es la permitividad dieléctrica y μ la permeabilidad magnética.

Si se aplica ahora, el rotacional en ambos lados de la Ec. (1.3) se tiene que

$$\nabla \times (\nabla \times \mathbf{E}) = -\frac{\partial}{\partial t} \nabla \times \mathbf{B}. \quad (1.5)$$

Utilizando la identidad vectorial

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}, \quad (1.6)$$

y tomando en cuenta que el medio de propagación es homogéneo (ϵ constante), sustituimos las Ec. (1.4) y (1.5) en (1.6) para obtener

$$\nabla^2 \mathbf{E} = \mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2}. \quad (1.7)$$

Algo similar ocurre con el campo magnético, es decir

$$\nabla^2 \mathbf{B} = \mu\epsilon \frac{\partial^2 \mathbf{B}}{\partial t^2}. \quad (1.8)$$

Ambas ecuaciones describen el comportamiento de una onda electromagnética en un medio lineal, homogéneo, isótropo y no dispersivo, cuya velocidad de propagación está dada por

$$v = \frac{1}{\sqrt{\mu\epsilon}}. \quad (1.9)$$

Ahora bien, dado que la velocidad de la luz en el vacío es

$$c = \frac{1}{\sqrt{\mu_0\epsilon_0}}, \quad (1.10)$$

podemos obtener el índice de refracción del medio dividiendo la velocidad de la luz entre la velocidad de propagación de la onda

$$n = \frac{c}{v} = \frac{\sqrt{\mu\epsilon}}{\sqrt{\mu_0\epsilon_0}}. \quad (1.11)$$

Sustituyendo esta última ecuación en las Ecs. (1.7) y (1.8) obtenemos

$$\nabla^2 \mathbf{E} - \frac{n^2}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0, \quad (1.12)$$

$$\nabla^2 \mathbf{B} - \frac{n^2}{c^2} \frac{\partial^2 \mathbf{B}}{\partial t^2} = 0. \quad (1.13)$$

Dado que las Ecs. (1.9) y (1.10) obedecen la ecuación de onda vectorial, cada una de las componentes vectoriales de dichas ecuaciones obedecen una ecuación de onda escalar, por lo que es posible sintetizar todas estas ecuaciones en una única ecuación de onda escalar [12]

$$\nabla^2 u_i(P, t) = \frac{n^2}{c^2} \frac{\partial^2 u_i(P, t)}{\partial t^2}, \quad (1.14)$$

donde $P = (x, y, z)$ y el subíndice i denota cualquier componente escalar del campo y resalta la dependencia espacio-temporal de las mismas.

En particular, para una onda monocromática, se tiene que

$$u(P, t) = A(P) \cos[2\pi\nu t + \phi(P)], \quad (1.15)$$

con $A(P)$ la amplitud de la onda, $\phi(P)$ su fase y ν la frecuencia. Utilizando notación compleja, la Ec. (1.15) puede escribirse como

$$u(P, t) = \text{Re}\{U(P) \exp(-i2\pi\nu t)\}, \quad (1.16)$$

donde $U(P)$ es una función compleja de la posición dada por

$$U(P) = A(P) \exp[-i\phi(P)]. \quad (1.17)$$

Sustituyendo la Ec. (1.16) en la Ec. (1.14) obtenemos

$$(\nabla^2 + k^2)U = 0, \quad (1.18)$$

conocida como ecuación de Helmholtz, la cual es independiente del tiempo y donde $k = 2\pi/\lambda$ es el número de onda, con λ la longitud de onda.

1.1.2. Principio de Huygens-Fresnel

Huygens consideró que cada punto en un frente de onda puede verse como una fuente de trenes de onda esféricos secundarios de manera que, el progreso del frente de onda puede ser determinado. Sin embargo, ésta técnica no toma en cuenta la mayoría de los trenes de onda secundarios y en consecuencia, el principio de Huygens no puede explicar los detalles del proceso de difracción. Fresnel fue quien dio solución a dicho problema con su concepto de interferencia. El principio de Huygens-Fresnel establece que [13] *cada punto sin obstrucción de un frente de onda, en un instante de tiempo determinado, sirve como fuente de trenes de onda secundarios esféricos (de la misma frecuencia que la onda primaria). La amplitud del campo óptico en cualquier punto más allá es la superposición de todos estos trenes de onda (considerando sus amplitudes y fases relativas).* El principio de Huygens-Fresnel presenta algunas limitaciones y fue Kirchhoff quien desarrolló una teoría más rigurosa que dio credibilidad a la suposición de Fresnel planteada anteriormente y que se basó directamente en el teorema integral de Helmholtz y Kirchhoff, la condición de radiación de Raileigh Sommerfeld, las condiciones de frontera de Kirchhoff y la solución de la ecuación diferencial de onda, dando como resultado la fórmula de difracción de Fresnel-Kirchhoff. Esta ecuación puede reescribirse de la siguiente manera

$$U(P_0) = \int \int_{\Sigma} U'(P_1) \frac{\exp(ikr_1)}{r_1} ds, \quad (1.19)$$

en donde Σ representa la superficie subtendida por la abertura,

$$U'(P_1) = \frac{1}{i\lambda} \left[\frac{A \exp(ikr_2)}{r_2} \right] \left[\frac{\cos(\mathbf{n}, \mathbf{r}_1) - \cos(\mathbf{n}, \mathbf{r}_2)}{2} \right], \quad (1.20)$$

r_1 es la distancia desde la apertura hasta el punto de observación, r_2 es la distancia de la fuente de iluminación a la apertura, \mathbf{n} es un vector normal a la apertura, λ es la longitud de onda de iluminación y ds es un elemento diferencial de superficie sobre la abertura.

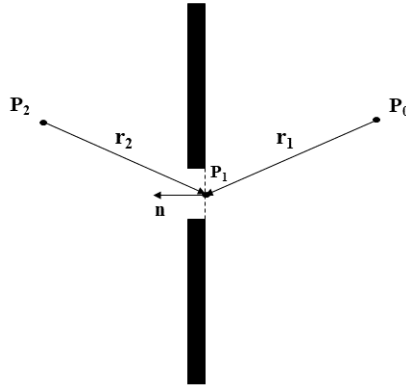


Figura 1.2: Iluminación de una abertura por una onda esférica proveniente de P_2 a una distancia r_2 de P_1 .

La Ec. (1.19) puede interpretarse como una superposición de ondas esféricas secundarias, que se originan en fuentes puntuales secundarias sobre el plano de la abertura. Las fuentes secundarias tienen ciertas amplitudes y fases descritas por $U'(P_1)$ relacionadas con el frente de onda de iluminación y los ángulos de iluminación y observación. La teoría de Kirchhoff es válida en el caso de longitudes de onda suficientemente pequeñas, es decir, cuando el tamaño de la abertura es grande en comparación con λ .

1.1.3. Formulación de difracción de Rayleigh-Sommerfeld

Aunque la teoría de Kirchhoff ha mostrado resultados precisos en la práctica, existen inconsistencias dentro de la teoría que motivaron la búsqueda de un desarrollo matemático más satisfactorio. La dificultad surge debido a que lo que se busca es la solución de una ecuación diferencial en derivadas parciales que satisfaga las condiciones de contorno impuestas tanto en el campo como en su derivada normal. En particular, es un conocido teorema de la teoría del potencial que si una función potencial de dos dimensiones y su derivada normal desaparecen juntas a lo largo de cualquier segmento de curva finita, entonces esa función potencial debe desaparecer por todo el plano. Del mismo modo, si una solución de la ecuación de onda tridimensional se desaparece en cualquier elemento de superficie finita, debe desaparecer en todo el espacio [14]. Así, las dos condiciones de contorno de Kirchhoff, implican que el campo es cero en todas partes detrás de la abertura, un resultado que contradice la situación física conocida. Una indicación adicional de estas inconsistencias, es el hecho de que la fórmula de difracción de Fresnel-Kirchhoff se puede demostrar que falla al reproducir las condiciones de contorno, asumiendo que el punto de observación se aproxima a la pantalla o la abertura [15].

Las inconsistencias de la teoría de Kirchhoff fueron eliminadas por Sommerfeld, eliminando la necesidad de imponer valores límite tanto en la perturbación y su derivada normal simultáneamente. Como resultado se llega a una expresión ligeramente diferente de la de Kirchhoff para el cálculo de la amplitud del campo óptico. Ésta es la llamada *fórmula de difracción de Rayleigh-Sommerfeld* [11].

$$U(P_0) = \frac{A}{i\lambda} \int \int_{\Sigma} \frac{\exp[ik(r_2 + r_1)]}{r_2 r_1} \cos(\mathbf{n}, \mathbf{r}_1) ds. \quad (1.21)$$

Este tratamiento es más consistente, pero las predicciones son muy similares a las que se obtienen al aplicar la solución de Kirchhoff.

1.1.4. Aproximación de Fresnel

A pesar de que la integral de Rayleigh-Sommerfeld permite obtener el campo difractado incluso cuando el plano de observación es cercano al objeto difractor, no puede ser resuelta de manera analítica, esto significa que deben realizarse aproximaciones dependiendo del tamaño de la abertura, la distancia al plano de observación y la región observada en el plano de salida.

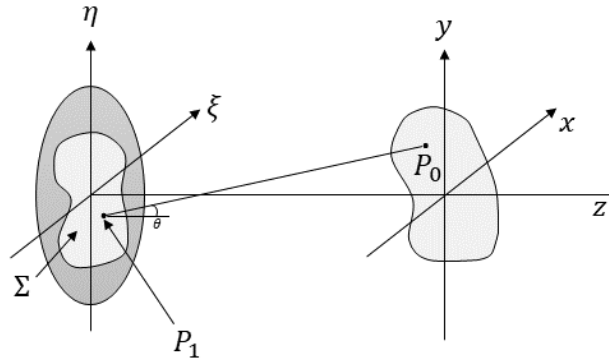


Figura 1.3: Geometría del fenómeno de difracción

Antes de introducir dichas aproximaciones al principio de Huygens-Fresnel, resulta útil expresarlo en coordenadas rectangulares. Tal como se observa en la Fig. (1.3) asumimos que la abertura de difracción se encuentra en el plano (ξ, η) y es iluminada en la dirección positiva de z . Se busca

calcular el campo a través del plano (x, y) , paralelo al plano (ξ, η) situado a una distancia z . De esta manera el, el principio de Huygens-Fresnel queda como

$$U(P_0) = \frac{1}{i\lambda} \int \int_{\Sigma} U'(P_1) \frac{\exp(ikr_1)}{r_1} \cos \theta ds, \quad (1.22)$$

con θ el ángulo entre los vectores \mathbf{n} y \mathbf{r}_1 . El término $\cos \theta$ está dado por

$$\cos \theta = \frac{z}{r_1}, \quad (1.23)$$

por lo que el principio de Huygens-Fresnel puede escribirse como

$$U(x, y) = \frac{z}{i\lambda} \int \int_{\Sigma} U(\xi, \eta) \frac{\exp(ikr_1)}{r_1^2} d\xi d\eta. \quad (1.24)$$

en donde

$$r_1 = \sqrt{z^2 + (x - \xi)^2 + (y - \eta)^2}. \quad (1.25)$$

Factorizando z de la Ec. (1.25) se tiene que

$$r_1 = z \sqrt{1 + \left(\frac{x - \xi}{z}\right)^2 + \left(\frac{y - \eta}{z}\right)^2}. \quad (1.26)$$

Utilizando la expansión binomial

$$\sqrt{1 + b} = 1 + \frac{1}{2}b + \frac{1}{8}b^2 + \dots \quad (1.27)$$

la Ec. (1.26) puede escribirse como

$$r_1 \approx z \left[1 + \frac{1}{2} \left(\frac{x - \xi}{z}\right)^2 + \frac{1}{2} \left(\frac{y - \eta}{z}\right)^2 \right]. \quad (1.28)$$

Sustituyendo la Ec. (1.27) en la Ec. (1.24) se tiene que

$$U(x, y) = \frac{\exp(ikz)}{i\lambda z} \iint_{-\infty}^{\infty} U(\xi, \eta) \exp \left\{ \frac{ik}{2z} [(x - \xi)^2 + (y - \eta)^2] \right\} d\xi d\eta. \quad (1.29)$$

A la ec.(1.29), se le conoce como la integral de difracción de Fresnel o de campo cercano, en la aproximación paraxial. Esta integral representa una superposición de ondas esféricas en la aproximación paraxial. La amplitud de estas ondas, por una parte, es proporcional a la amplitud compleja del campo que incide en la abertura, y por otra parte, es proporcional al factor $1/\lambda$. Así mismo, estas ondas tienen un corrimiento de fase de $\pi/2$, que es indicado en el factor de fase $1/i$. Se puede observar que la integral de difracción de Fresnel, es una solución a la ecuación paraxial de ondas esféricas. A partir de esta expresión es posible hacer nuevas aproximaciones para ángulos pequeños, esto es, si $f_x = \frac{\alpha}{\lambda}$ es la frecuencia espacial en la dirección x , con α el coseno director del vector de onda en la dirección x entonces

$$|\lambda f_x| \ll 1, \quad (1.30)$$

es conocida como aproximación de Fresnel o de ángulos pequeños, la cual es equivalente a la aproximación paraxial. Si adicionalmente consideramos que

$$z \gg \frac{\pi(\xi^2 + \eta^2)}{\lambda}, \quad (1.31)$$

entonces estamos en el régimen o aproximación de Fraunhofer también conocido como aproximación de campo lejano. En ambos casos resulta interesante que el campo difractado es expresado en términos de una transformada de Fourier. Este hecho es fundamental para la evaluación no solo del campo sino también para implementar el esquema de recuperación de fase del campo. Esto se debe gracias a que la transformada de Fourier es un numero complejo y por tanto posee modulo y fase. En la siguiente sección veremos como estas cantidades pueden tratarse en términos de un formalismo matemático basado en proyecciones en espacios convexos y no convexos. Este hecho permite implementar un algoritmo iterativo basado en la sustitución del modulo calculado por el modulo medido para obtener la fase asociada a dicho modulo.

1.2. Medición de un campo óptico como una proyección

1.2.1. Detección de amplitud y fase

Para obtener una descripción completa de un frente de onda complejo es necesario conocer tanto la amplitud de la onda así como su fase. De acuerdo a la teoría escalar de difracción vista en la sección anterior, el campo luminoso en el plano del objeto, puede escribirse como:

$$f(x, y) = |f(x, y)|e^{i\phi(x, y)}, \quad (1.32)$$

donde $|f(x, y)|$ es la amplitud del campo mientras que $\phi(x, y)$ representa la fase. Dada la naturaleza de un detector, estamos limitados a registrar únicamente la intensidad de la onda difractada, proporcional a la raíz cuadrada de la intensidad medida I , esto es, $|f(x, y)| \sim \sqrt{I}$. Sin embargo, la información de la fase se pierde. Esto es conocido como el problema de fase en óptica.

1.2.2. Recuperación de fase y factibilidad

Consideramos un espacio funcional \mathcal{L} . De manera general, el problema de recuperación de una señal consiste en estimar la forma original de una señal x en este espacio sujeta a las restricciones impuestas por las mediciones de señales relacionadas físicamente por una transformación lineal e información a priori. En el caso de recuperación de fase, estas mediciones son las intensidades medidas en el plano de Fourier, es decir, el módulo m de la Transformada de Fourier \mathcal{F} de la señal x [10]

$$|\mathcal{F}| = m. \quad (1.33)$$

Un espacio de señales general que modela apropiadamente la física involucrada es el espacio de Hilbert \mathcal{H} . La medición m es una función no negativa que es proporcional al módulo de la transformada de Fourier, mientras que, $x \in \mathcal{H}$ es una función cuadrado-integrable que mapea una función continua $t \in \mathbb{R}^N$ a un número complejo $x(t) \in \mathbb{C}$. El conjunto de señales que satisfacen la restricción en el dominio de Fourier es

$$M = \{y \in \mathcal{H} : |\mathcal{F}(y)| = m\}. \quad (1.34)$$

En la recuperación de fase, M es el conjunto de señales cuyo módulo de Fourier concuerda con algunos datos medidos. En este caso, m es una medida de la magnitud de la transformada de Fourier de la señal original x definida por $m : \mathbb{R}^N \rightarrow \mathbb{R}$; por tanto, m es una función par no negativa. La proyección de una señal x sobre el conjunto M puede ser calculada reemplazando la magnitud de la Transformada de Fourier de la señal x por los datos de la magnitud conocida m y aplicando la Transformada de Fourier inversa al resultado, es decir

$$P_M(x) = \mathcal{F}^{-1}(G), \quad (1.35)$$

en donde \mathcal{F}^{-1} es la Transformada Inversa de Fourier y G es una selección particular de la proyección del dominio de Fourier, definida por

$$G = \mathcal{F}(y_0(\omega)) = \begin{cases} m(\omega) \frac{\mathcal{F}(x(\omega))}{|\mathcal{F}(y(\omega))|}, & \mathcal{F}(x(\omega)) \neq 0, \\ m(\omega), & \text{de otra forma.} \end{cases} \quad (1.36)$$

Por otra parte, considerando que el soporte de la señal x está contenido en algún conjunto $D \in \mathbb{R}^N$. Denotando la función característica de un conjunto $E \in \mathbb{R}^N$ como 1_E y como E^C a su complemento, esta restricción en el dominio del objeto confina x al conjunto

$$S = \{y \in \mathcal{H} : y \cdot 1_{D^C} = 0\}. \quad (1.37)$$

De esta manera, el problema de recuperación de fase puede plantearse como el problema de encontrar una señal $x \in \mathcal{H}$ que satisfaga ambas restricciones, es decir, encontrar una señal que cumpla con [16]

$$x \in S \cap M. \quad (1.38)$$

Se trata de encontrar un punto de intersección entre ambos conjuntos de restricciones, es decir un problema de estimación de la teoría de conjuntos. En matemáticas, los problemas de este tipo se denominan problemas de factibilidad. La importancia del enfoque de la teoría de conjuntos es permitir la incorporación de una amplia gama de información estadística y no estadística en la definición de una solución.

La recuperación de señales sigue siendo hoy en día el campo de aplicación más activo de la estimación de la teoría de conjuntos debido a dos factores importantes. En primer lugar los problemas de recuperación de señales suelen ir acompañados de una gran cantidad de información cualitativa sobre la señal original que no se expresa fácilmente en términos puramente estadísticos, que es la única forma que pueden aprovechar los métodos de estimación convencionales. El segundo factor es que, en la mayoría de los casos, un observador humano es quien juzga la calidad de la señal recuperada y dado que el ojo humano no es sensible a las medidas matemáticas estándar, la importancia de una recuperación óptima se reduce significativamente.

1.3. Proyecciones en espacios convexos y no convexos

1.3.1. Distancia, proyecciones y proyectores

1. Distancia

Suponemos que \mathcal{H} es el espacio general de Hilbert con producto interno $\langle \cdot, \cdot \rangle$ y norma $\| \cdot \| : x \mapsto \sqrt{\langle x, x \rangle}$. [17]

Definición 1. Distancia. Si Y es un conjunto en \mathcal{H} y $x \in \mathcal{H}$, la distancia de x a Y es

$$d(x, Y) = \inf\{d(x, y) : y \in Y\}, \quad (1.39)$$

en donde $d(x, y)$ es la distancia de x a un punto $y \in \mathcal{H}$ definida como

$$d(x, y) = \|x - y\|. \quad (1.40)$$

2. Operadores de proyección

Definición 2. Operador de proyección. Sea Y un conjunto en \mathcal{H} . Si $x \in \mathcal{H}$, entonces el conjunto de puntos en Y más cercanos a x , es decir

$$\{y \in Y : d(x, y) = d(x, Y)\}, \quad (1.41)$$

es llamado la proyección de x en Y y denotado como $\Pi_Y(x)$. El operador Π_Y es llamado operador de proyección.[18] En el caso de procesamiento de señales, si Y contiene las señales que satisfacen una determinada propiedad, entonces las señales en Π_Y son las señales más cercanas a x que satisfacen esta propiedad. Es importante notar que la salida de un proyector Π_Y forma subconjuntos de Y . Estos pueden estar vacíos, reducirse a un solo elemento o tener más de un elemento:

3. Proyecciones para el problema de recuperación de fase

En el caso del problema de recuperación de fase, el espacio que modela mejor la física involucrada es el espacio complejo de Hilbert, denotado por:

$$\mathcal{L} = L^2[\mathbb{R}^2, \mathbb{C}]. \tag{1.42}$$

El enfoque más común para resolver el problema de recuperación de fase es hacer cumplir las restricciones del dominio de objeto conocido y del dominio de Fourier de alguna manera alterna . Por lo tanto, dada una señal x , la restricción de soporte se aplica naturalmente estableciendo x igual a cero fuera del dominio D dado a través de la transformación $x \mapsto x \cdot 1_D$. Esta simple operación es en realidad una proyección.

Restricción de soporte: Suponga que D es un conjunto medible en \mathbb{R}^N y fije $x \in \mathcal{L}$. Entonces la proyección de x sobre el conjunto S (Eq. 1.14, sección 1.2) es:

$$P_S(x) = x \cdot 1_D. \tag{1.43}$$

La misma observación es válida para la restricción del módulo de imagen. A continuación, se describen los enfoques para hacerla cumplir; nuevamente, estas operaciones resultan ser proyecciones.

Restricción de módulo de imagen: Sea m una función no negativa en \mathcal{L} , y fije $x \in \mathcal{L}$. Entonces $y \in \mathcal{L}$ pertenece a la proyección $\Pi_M(x)$ de x sobre el conjunto M de la ecuación 1.13 si y solo si satisface que

$$\mathcal{F}(y(\omega)) = \begin{cases} m(\omega) \frac{\mathcal{F}(x(\omega))}{|\mathcal{F}(y(\omega))|}, & \mathcal{F}(x(\omega)) \neq 0, \\ m(\omega) \exp[i\phi(\omega)], & \text{de otra forma,} \end{cases} \tag{1.44}$$

para alguna función medible $\phi : \mathbb{R}^N \rightarrow \mathbb{R}$.

1.3.2. Convexidad y cerradura

En lo que sigue, asumimos que \mathcal{H} es el espacio de Hilbert real con producto interno $\langle \cdot, \cdot \rangle$ y norma inducida $\| \cdot \|$

Definición 3. Subespacio vectorial y conjunto convexo. Suponga que C es un conjunto no vacío de \mathcal{H} . Entonces C es

i) subespacio vectorial si el vector cero esta contenido en C y si se cumple además que

$$\lambda c_1 + (1 - \lambda)c_2 \in C, \tag{1.45}$$

para cualesquiera $c_1, c_2 \in C$ y $\lambda \in \mathbb{R}$.

ii) conjunto convexo si se cumple que

$$\lambda c_1 + (1 - \lambda)c_2 \in C, \tag{1.46}$$

para cualesquiera $c_1, c_2 \in C$ y $\lambda \in [0, 1]$. Es decir, si el segmento de recta que une dos puntos cualesquiera en C se encuentra completamente en C . [19]

Definición 4. Conjunto cerrado. Suponga que C es un conjunto en \mathcal{H} . Entonces C es un conjunto cerrado si, siempre que (c_n) es una sucesión en C que converge a algún punto $c \in \mathcal{H}$, el punto límite c pertenece necesariamente a C .

En los algoritmos descritos más adelante, la importancia de la cerradura radica en que a menudo se quiere que el límite de una sucesión herede buenas propiedades de los términos de dicha sucesión.

La cerradura es un condición necesaria para garantizar la existencia de proyecciones. De hecho, un punto x en la cercanía de C pero no en C no tiene proyección sobre C . Para espacios de dimensión finita, la cercanía también es suficiente para garantizar la existencia de proyecciones; sin embargo, para espacios de dimensión infinita, esto ya no se cumple. [20,21]

Capítulo 2

Métodos iterativos de Fourier basados en el algoritmo de Gerchberg-Saxton

2.1. Transformada y transformada rápida de Fourier

Consideramos una función bidimensional $f(x, y)$ que cumple con las siguientes condiciones [11]:

- 1.- f es absolutamente integrable en el plano infinito (x, y) .
- 2.- f tiene solo un número finito de discontinuidades y un número finito de máximos.
- 3.- f no tiene discontinuidades infinitas.

Se define la Transformada de Fourier de la función f , representada por $\mathcal{F}\{f\}$, como

$$G(p, q) = \mathcal{F}\{f(x, y)\} = \iint_{-\infty}^{\infty} f(x, y)e^{-i2\pi(px+qy)} dx dy. \quad (2.1)$$

De manera similar, la Transformada de Fourier Inversa de la función $G(p, q)$, representada por $\mathcal{F}^{-1}\{G\}$, está definida como

$$f(x, y) = \mathcal{F}^{-1}\{G(p, q)\} = \iint_{-\infty}^{\infty} G(p, q)e^{i2\pi(px+qy)} dp dq. \quad (2.2)$$

La transformada es en sí misma una función de valor complejo de las variables p y q , a las que generalmente nos referimos como frecuencias y $G(p, q)$ es el espectro de Fourier.

2.1.1. Propiedades

Algunas de las propiedades que posee la Transformada de Fourier son:

1. **Linealidad:** Si $f_1(x, y) \xrightarrow{\mathcal{F}} G_1(p, q)$, $f_2(x, y) \xrightarrow{\mathcal{F}} G_2(p, q)$ y $\alpha, \beta \in \mathbb{R}$ entonces:

$$[\alpha f_1(x, y) + \beta f_2(x, y)] \xrightarrow{\mathcal{F}} [\alpha G_1(p, q) + \beta G_2(p, q)]. \quad (2.3)$$

2. **Cambio de escala:** Si $f(x, y) \xrightarrow{\mathcal{F}} G(p, q)$ y $\alpha, \beta \in \mathbb{R}$ entonces

$$f(\alpha x, \beta y) \xrightarrow{\mathcal{F}} \frac{1}{|\alpha\beta|} G\left(\frac{p}{\alpha}, \frac{q}{\beta}\right), \quad (2.4)$$

es decir, un alargamiento en el dominio del espacio da como resultado una contracción en el dominio de las frecuencias multiplicada por un factor constante.

3. Desplazamiento:

En el dominio del espacio: Si $f(x, y) \xrightarrow{\mathcal{F}} G(p, q)$ y $a, b \in \mathbb{R}$ entonces

$$f(x - a, y - b) \xrightarrow{\mathcal{F}} G(p, q)e^{-i2\pi(ap+bq)}. \quad (2.5)$$

En el dominio de frecuencias: Si $f(x, y) \xrightarrow{\mathcal{F}} G(p, q)$ y $a, b \in \mathbb{R}$ entonces

$$f(x, y)e^{i2\pi(ap+bq)} \xrightarrow{\mathcal{F}} G(p - a, q - b). \quad (2.6)$$

4. Teorema de Parseval: Si $\mathcal{F}\{f(x, y)\} = G(p, q)$ entonces

$$\iint_{-\infty}^{\infty} |f(x, y)|^2 dx dy = \iint_{-\infty}^{\infty} |G(p, q)|^2 dp dq. \quad (2.7)$$

La integral de lado izquierdo puede interpretarse como la energía contenida en el frente de onda $f(x, y)$.

5. Teorema Integral de la Transformada de Fourier: Si $\mathcal{F}\{f(x, y)\} = G(p, q)$ entonces

$$\mathcal{F}\mathcal{F}^{-1}\{f(x, y)\} = \mathcal{F}^{-1}\mathcal{F}\{f(x, y)\} = f(x, y), \quad (2.8)$$

válido para los puntos en los cuales $f(x, y)$ es continua.

2.1.2. Discretización de la Transformada y Transformada Inversa de Fourier

Sabemos que en situaciones reales, tanto la señal de interés como su Transformada de Fourier son funciones continuas del espacio y de la frecuencia respectivamente. A la hora de medir dicha señal, es posible obtener únicamente un número finito de datos y todo el procesamiento posterior se realiza en computadoras digitales. Por ello en lugar de trabajar con funciones continuas se trabaja con aproximaciones discretas de $f(x, y)$ y $G(p, q)$ que están bien justificadas dada la resolución finita que deriva de las mediciones y, en general, del hecho de que todos los sistemas ópticos tienen límites de resolución. Para los algoritmos que se desarrollan en este trabajo, la discretización de las ecuaciones presentadas anteriormente se describe a continuación.

En el caso discreto, la función $f(x, y)$ está definida sobre un conjunto de puntos finitos y puede ser escrita como:

$$f(m\Delta x, n\Delta y) = f_{m,n}, \quad (2.9)$$

siendo Δx y Δy el intervalo de muestreo en las direcciones x y y respectivamente y m y n representan los índices de los valores de las muestras. Si asumimos que el espacio muestral está compuesto de N muestras en la dirección x y M muestras en la dirección y , m y n son definidas como [23]

$$m = -\frac{M}{2}, \dots, \frac{M}{2} - 1, \quad n = -\frac{N}{2}, \dots, \frac{N}{2} - 1. \quad (2.10)$$

La Transformada de Fourier Discreta en dos dimensiones es entonces [24]

$$F_{p,q} = \sum_{m,n} f_{m,n} e^{-i2\pi\left(\frac{mp}{N}, \frac{nq}{M}\right)}. \quad (2.11)$$

Similarmente, la Transformada de Fourier Inversa es

$$f_{n,m} = \frac{1}{NM} \sum_{p,q} F_{p,q} e^{i2\pi\left(\frac{np}{N}, \frac{mq}{M}\right)}, \quad (2.12)$$

donde p y q son números enteros que indexan múltiplos de los intervalos de muestreo de frecuencia, definidos de manera similar que las ecuaciones (2.10). Es posible deducir a partir de la ecuación (2.10), que la Transformada de Fourier Discreta (TFD) en dos dimensiones está relacionada con la Transformada de Fourier Discreta en una dimensión

$$F_{p,q} = \sum_n \left(\sum_m f_{n,m} e^{-i2\pi\left(\frac{np}{N}\right)} \right) e^{-i2\pi\left(\frac{mq}{M}\right)},$$

$$F_{p,q} = \sum_n TFD_n \{f_{n,m}\} e^{-i2\pi\left(\frac{mq}{M}\right)}. \quad (2.13)$$

Observamos entonces que, para obtener la TFD en dos dimensiones del arreglo bidimensional $f_{n,m}$ basta con tomar la TFD en una dimensión sobre cada fila de la matriz de tamaño $N \times M$ y después sobre cada columna del resultado. Lo mismo sucede con la TFD inversa.

2.1.3. Transformada rápida de Fourier en MATLAB (FFT)

Para las simulaciones numéricas realizadas en este trabajo se utilizó MATLAB. En este programa resulta sencillo implementar los algoritmos iterativos propuestos en las siguientes secciones a través de bucles o ciclos en donde se evalúa alguna expresión y se repite la ejecución de un grupo de instrucciones hasta cumplir una condición que haga verdadera a la expresión.

MATLAB cuenta con una instrucción que calcula la Transformada de Fourier Discreta usando un algoritmo de transformada rápida de Fourier (FFT, *Fast Fourier Transform*) en una y dos dimensiones. Los comandos utilizados para calcular la TFD son `fft` y `fft2` para una y dos dimensiones respectivamente. Algo similar ocurre con la Transformada de Fourier Discreta Inversa en donde los comandos utilizados son `ifft` e `ifft2` para una y dos dimensiones. Dado que las señales que se trabajarán aquí están definidas en un espacio bidimensional, los comandos que se utilizarán más adelante serán `fft2` e `ifft2`.

Es importante notar que si se implementa el algoritmo de la FFT de manera directa, los resultados que se muestran en la pantalla aparecen desordenados debido a la forma en la que MATLAB asigna los valores de frecuencia en cada arreglo de vectores o matrices. Para visualizar de manera correcta los resultados obtenidos es necesario utilizar los comandos `fftshift` e `ifftshift`. Ambos comandos reorganizan una transformada de Fourier (directa e inversa respectivamente) desplazando la componente de frecuencia cero al centro de la matriz.

2.2. Algoritmo de Gerchberg-Saxton

En 1972 R.W Gerchberg y W. O Saxton, a través de su artículo *A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures*, propusieron un algoritmo como solución al problema de recuperación de fase para una función de onda cuya intensidad en los planos objeto y de Fourier se conocen. Este algoritmo es un ejemplo clásico de técnicas de optimización conocidas como proyecciones alternantes. [25]. El algoritmo básico consiste en la aplicación sistemática de la Transformada de Fourier (TF) y la Transformada de Fourier Inversa (TFI). Los datos de entrada son las amplitudes en el plano imagen y el plano de difracción, proporcionales a las raíces cuadradas de las intensidades medidas por los detectores de luz.

Para iniciar el algoritmo, se elige una fase aleatoria en un intervalo de 0 a 2π , sin embargo, puede utilizarse también una fase que ha sido predicha con algún modelo matemático. Ésta fase será la primera aproximación de la fase buscada y será modificada con cada iteración del algoritmo. El paso siguiente es multiplicar el término de fase inicial con la amplitud muestreada en el plano del objeto para posteriormente obtener su TF. En este punto se aplica la restricción en el plano de difracción sustituyendo el módulo proveniente de la TF por el módulo medido, conservando la fase. Esta nueva función es llevada nuevamente al plano objeto empleando la TFI para luego ser

modificada aplicando ahora la restricción en este plano y así formar una nueva estimación de la función compleja en el plano objeto y repetir el proceso.

El algoritmo de Gerchberg-Saxton (GS) también ha sido generalizado con el fin de obtener proyecciones bien definidas y de manera que sea computacionalmente eficiente. Una forma de lograr esta eficiencia y de que el algoritmo trabaje bien es imponiendo la restricción de un soporte limitado del objeto, es decir, sabiendo que algunas partes de la señal son cero. Las imposiciones de las restricciones que se dan tanto en el plano del objeto como en el plano de Fourier se hacen a través de proyecciones, las cuales garantizan la convergencia del algoritmo. Dicha convergencia puede verse afectada al llegar a un punto estacionario que no es precisamente la solución o bien, que el algoritmo se encuentre en un bucle sin fin.

Por último, el algoritmo GS ha sido aplicado con éxito en la recuperación de fase a partir de mediciones en áreas como astronomía, cristalografía de rayos x, además de óptica.

2.3. Estructura de algoritmos iterativos para recuperación de fase

Los métodos iterativos para el problema de recuperación de fase que se utilizan hoy en día están basados esencialmente en el algoritmo de Gerchberg-Saxton. Se trata de procedimientos iterativos en donde se aplican la TF y la TFI entre los espacios del dominio del objeto (donde se aplican los conocimientos previos sobre el objeto) y el dominio de Fourier (donde se aplican los datos del módulo de Fourier). Cada iteración de éste algoritmo (descrito en la sección anterior) consiste de los siguientes pasos:

1. Transformada de Fourier del objeto estimado.
2. Reemplazar el módulo de la TF calculada con el módulo de Fourier medido para formar una nueva estimación de la TF.
3. Transformada de Fourier Inversa de la estimación de la TF.
4. Reemplazar el módulo del objeto calculado con el módulo del objeto medido para formar una nueva estimación del objeto. Escrito matemáticamente, los pasos anteriores para la k -ésima iteración son[26]:

$$G_k(p, q) = |G_k(p, q)| \exp[i\phi_k(p, q)] = \mathcal{F}[g_k(x, y)], \quad (2.14)$$

$$G'_k(p, q) = |F(p, q)| \exp[i\phi_k(p, q)], \quad (2.15)$$

$$g'_k(x, y) = |g'_k(x, y)| \exp[i\theta'_k(x, y)] = \mathcal{F}^{-1}[G'_k(p, q)], \quad (2.16)$$

$$g_{k+1}(x, y) = |f(x, y)| \exp[i\theta_{k+1}(x, y)] = |f(x, y)| \exp[i\theta'_k(x, y)], \quad (2.17)$$

donde $|F|$ es el módulo de Fourier medido y $|f|$ es el modulo en el dominio del objeto. Posteriormente James R. Fienup (1982) logró mejorar la velocidad de convergencia del algoritmo mediante los algoritmos llamados de entrada y salida: el algoritmo de reducción de error (ER), el algoritmo básico de entrada-salida (BIO) y el algoritmo híbrido de entrada-salida (HIO). En ellos se requiere una única medida de intensidad en el dominio del espacio de Fourier para recuperar la fase deseada.

ER (Error Reduction): El algoritmo de reducción de error es una generalización del algoritmo GS, es el más sencillo y el primero en ser considerado por Fienup. La operaciones son las mismas que en el algoritmo GS pero difiere en la actualización de la función en el dominio del objeto (Ec. 2.16), es decir, la siguiente iteración estará dada por:[10]

$$g_{k+1}(x, y) = \begin{cases} g'_k(x, y), & (x, y) \notin \gamma, \\ 0, & (x, y) \in \gamma. \end{cases} \quad (2.18)$$

en donde γ es el conjunto de puntos en los cuales $f'(x, y)$ viola las restricciones en el dominio del objeto (no negatividad y de soporte).

BIO (*Basic Input-Output*): Con el fin de solucionar el problema de la lenta convergencia del algoritmo ER, Fienup propuso el algoritmo básico de entrada-salida. Si los pasos (2.14 - 2.17) en conjunto se piensan como un sistema no lineal cuya entrada es $f(x, y)$ y su salida es $f'(x, y)$, podemos notar que la salida $f'(x, y)$ es siempre una señal que tiene una transformada de Fourier que satisface las restricciones del dominio de Fourier. Por lo tanto, si la salida también satisface las restricciones del dominio del objeto, es una solución del problema. La entrada $f(x, y)$ se puede considerar como una función impulsora para la siguiente salida y no necesita satisfacer las restricciones del dominio del objeto. Siguiendo esta idea, se puede decir que un pequeño cambio de la entrada da como resultado un cambio de la salida en la misma dirección que el cambio de la entrada. En otras palabras, para un pequeño cambio de la entrada, el cambio de la salida es una constante α multiplicada por el cambio de la entrada. Si se desea un cambio Δx a la salida, una elección lógica del cambio de la entrada para lograr ese cambio de la salida sería $\beta\Delta x$, con β una constante idealmente igual α^{-1} . En el problema planteado, para una sola medición de intensidad, nuevamente las ecuaciones (2.3-2.15) permanecen igual y el cambio deseado a la salida es:

$$\Delta g_k(x, y) = \begin{cases} 0, & (x, y) \notin \gamma, \\ -g'_k(x, y), & (x, y) \in \gamma. \end{cases} \quad (2.19)$$

en donde γ es el conjunto de puntos en los cuales $f'(x, y)$ viola las restricciones en el dominio del objeto (no negativas y de soporte). Donde se satisfacen dichas restricciones, no hay un cambio en la salida; pero donde estas son violadas, el cambio deseado de la salida, para que satisfaga las restricciones del dominio del objeto, debe ser cero, por lo tanto, el cambio deseado es la salida negada en esos puntos. Así, una elección lógica para la siguiente entrada es:

$$\begin{aligned} g_{k+1}(x, y) &= g_k(x, y) + \beta\Delta g_k(x, y), \\ &= \begin{cases} g_k(x, y), & (x, y) \notin \gamma, \\ g_k(x, y) - \beta g'_k(x, y). & (x, y) \in \gamma. \end{cases} \end{aligned} \quad (2.20)$$

Debido a que la TF de g' ya satisface las restricciones del dominio de Fourier y por tanto no se ve afectada a medida que pasa por el sistema no lineal, si utilizamos una salida g' como entrada, la salida será ella misma. De esta manera, otra elección para la siguiente entrada es:

$$\begin{aligned} g_{k+1}(x, y) &= g'_k(x, y) + \beta\Delta g_k(x, y), \\ &= \begin{cases} g'_k(x, y), & (x, y) \notin \gamma, \\ g'_k(x, y) - \beta g'_k(x, y), & (x, y) \in \gamma. \end{cases} \end{aligned} \quad (2.21)$$

La ecuación (2.20) hace referencia al algoritmo de *salida-salida*. Note que si $\beta = 1$, el algoritmo se reduce al algoritmo ER.

HIO (*Híbrido Input-Output*): El tercer algoritmo propuesto por Fienup es una combinación de las ecuaciones (2.19) y (2.20). La siguiente entrada del algoritmo está dada por:

$$g_{k+1}(x, y) = \begin{cases} g'_k(x, y), & (x, y) \notin \gamma, \\ g_k(x, y) - \beta g'_k(x, y), & (x, y) \in \gamma. \end{cases} \quad (2.22)$$

El algoritmo híbrido de entrada-salida es el más utilizado actualmente debido a su simplicidad y generalmente ofrece una mejor velocidad de convergencia que los demás algoritmos.

Capítulo 3

Generación de un vórtice óptico usando técnicas de recuperación de fase

3.1. Vórtice óptico: definición y características principales

3.1.1. Momento angular orbital

El momento angular orbital (OAM) de un haz de luz está relacionado con la estructura espacial de su fase. El ejemplo más sencillo de un haz de luz que lleva OAM es uno con una fase en el plano transversal $f(r, \phi) = \exp(in\phi)$ en donde ϕ es la coordenada angular y n es un número entero, positivo o negativo. Estos haces poseen frentes de fase helicoidal con un número de hélices entrelazadas dependientes de la magnitud y signo de n . El vector de Poynting, que es en todo momento paralelo a la normal de la superficie de estos frentes de fase, tiene una componente azimutal alrededor del haz y, por lo tanto, un momento angular a lo largo del eje del haz.[27]

Un hecho sorprendente es que el OAM de un haz helicoidal al igual que el espín, esta en unidades de \hbar . Este hecho se deduce de un argumento geométrico simple: En un radio r , la inclinación del frente de fase, y por lo tanto del vector de Poynting, con respecto al eje del haz es simplemente $n\lambda/2\pi r$. Esto a su vez, establece la componente azimutal del momento lineal de la luz como $\hbar k_0 n \lambda / 2\pi r$ por fotón, el cual, cuando es multiplicado por el vector radio, da un momento angular $n\hbar$ por fotón [2].

3.1.2. Vórtice Óptico

Los vórtices ópticos están caracterizados principalmente por su perfil espacial anular (con un spot central oscuro) y un frente de onda en forma helicoidal (que se extiende a lo largo de la dirección de propagación del haz) capaz de transferir el momento angular orbital a partículas dieléctricas a una velocidad proporcional a su carga topológica, a diferencia de un haz Gaussiano convencional, cuyo frente de onda es plano. Las propiedades espaciales y de fase de un haz Gaussiano y un vórtice se muestran en la Fig. 3.1.2.

El término de fase de un vórtice óptico está caracterizado por la ecuación:

$$\exp(in\phi), \tag{3.1}$$

Generación de un vórtice óptico usando técnicas de recuperación de fase
 3.1 Vórtice óptico: definición y características principales

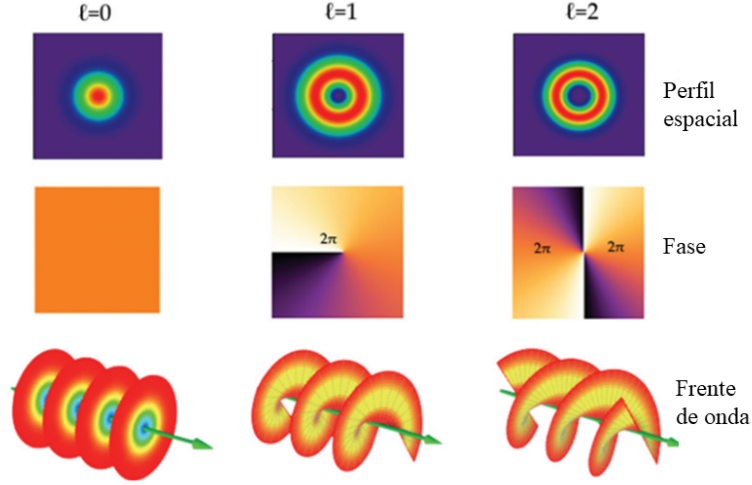


Figura 3.1: Comparación entre los perfiles espaciales y frentes de onda de un haz Gaussiano y un vórtice

en donde la carga topológica n es un número entero que representa el salto de fase de 2π del frente de onda alrededor del eje de propagación. Debido al frente de onda helicoidal, los vórtices poseen momento angular orbital (OAM), en donde cada fotón transporta un momento angular orbital de $n\hbar$, en donde \hbar es la constante de Plank reducida. En la singularidad, la fase está indeterminada y la polarización y la amplitud desaparecen por completo, lo que da como resultado un centro oscuro dentro de un conjunto de anillos concéntricos. Los vórtices ópticos ocurren típicamente en ausencia de condiciones especiales: son estructuras naturales. También se producen en luz estructurada, por ejemplo, en haces de luz cilíndricos con una fase uniformemente creciente alrededor del eje (como lo haces de Laguerre-Gauss). Alrededor del eje de dichos haces, la fase cambia en $2\pi n$ para el entero n , que define la carga asociada al vórtice.

Las propiedades que resultan ser más interesantes de los vórtices ópticos están relacionadas principalmente con la estructura de la fase y su carga topológica. Como se sabe, dicha estructura solo es posible ser representada en el espacio de la fase. Al aumentar el ángulo de fase de manera continua (en el sentido de las agujas del reloj) de 0 a $2\pi n$ y de nuevo al origen, el ángulo cero es ahora $2\pi n$ formando así una distribución de fase continua a lo largo de un bucle cerrado. El punto central del bucle cerrado donde no se puede definir la fase es la singularidad de fase mencionada anteriormente.

La carga topológica de un vórtice óptico con distribución de fase ϕ viene dada por [28]

$$n = \frac{1}{2\pi} \times \oint_C \nabla\phi(\mathbf{r})dr, \quad (3.2)$$

donde la integral es tomada a lo largo de una trayectoria C que rodea la singularidad. La magnitud de la carga topológica influye en el tamaño del spot obscuro central. La dirección de giro del frente de onda, ya sea en sentido horario o en sentido anti horario indica el signo de la carga. Para un signo positivo (+) el vórtice tiene un frente de onda que gira en el sentido de las agujas del reloj, mientras que para un signo negativo (-) el vórtice tiene un frente de onda que gira en sentido contrario al giro de las manecillas del reloj.

3.2. Esquemas de generación de vórtices ópticos: coherentes y parcialmente coherentes

Se ha hablado ya de las características físicas y la descripción matemática de un vórtice óptico. Se mencionó que un haz con vórtice tiene la característica de tener un frente de onda helicoidal que permite la transferencia de momento angular orbital y lo hace de manera proporcional a su carga topológica. En ésta sección se hablará de dos técnicas de generación de vórtices ópticos existentes en la actualidad, el primero completamente coherente y el siguiente con coherencia estructurada.

Generación de un vórtice óptico perfecto utilizando un LC-SLM

La idea de un vórtice óptico perfecto fue propuesta por Ostrovsky [29]. Su principal característica es que su radio no depende de la carga topológica, a diferencia de los vórtices ópticos convencionales. Se define un vórtice óptico perfecto como un haz con una distribución transversal de campo dada por:

$$g_\nu(\rho, \theta) \equiv \delta(\rho - \rho_0) \exp(i\nu\theta), \quad (3.3)$$

donde $\delta(\rho)$ es la función delta de Dirac, (ρ, θ) son las coordenadas polares en la sección transversal del haz, ρ_0 es el radio del vórtice y ν es la carga topológica del vórtice. La función anterior puede ser expandida en una serie de Bessel [30]

$$g(\rho) = \sum_{n=1}^{\infty} C_{\nu,n} J_\nu(\alpha_{\nu,n} \frac{\rho}{a}), \quad (3.4)$$

con

$$C_{\nu,n} = \frac{2}{a^2 [J_{\nu+1}(\alpha_{\nu,n})]^2} \int_0^a g(\rho) J_\nu(\alpha_{\nu,n} \frac{\rho}{a}) \rho d\rho, \quad (3.5)$$

y $J_\nu(\cdot)$ es la función de Bessel de primer tipo y ν -ésimo orden, $\alpha_{\nu,n}$ es la n -ésima raíz de la función $J_\nu(\cdot)$. Asumiendo que $a > \rho_0$ y sustituyendo la Ec.(3.3) en la Ec.(3.5) y utilizando además la propiedad de filtración de la función δ , la Ec. (3.4) representa al vórtice óptico como

$$g_\nu(\rho, \theta) \propto \text{circ}\left(\frac{\rho}{a}\right) \exp(i\nu\theta) \sum_{n=1}^{\infty} \frac{J_\nu(\alpha_{\nu,n} \rho_0/a)}{[J_{\nu+1}(\alpha_{\nu,n})]^2} J_\nu\left(\alpha_{\nu,n} \frac{\rho}{a}\right). \quad (3.6)$$

En la Fig ?? se muestra el sistema óptico utilizado para generar éste tipo de vórtice. El elemento principal de este sistema es un LC-SLM de transmisión situado entre dos polarizadores cuyas orientaciones de los ejes de transmisión se eligen de manera que el LC-SLM opera en el modo de modulación de solo fase. El campo óptico resultante puede ser observado en el plano focal de una lente delgada que realiza la transformada de Fourier.

Si se asume que el LC-SLM tiene una transmitancia conformada por un conjunto de anillos concéntricos con fase azimutal separados por regiones opacas donde la transmitancia es nula. Adicionalmente los anillos son lo suficientemente delgados, la transmitancia del LC-SLM puede ser representada por la función compleja

$$t_\nu(r, \phi) = \sum_{n=1}^N \alpha_{\nu,n} \beta_{\nu,n} \exp[i\nu(\phi - \phi_{\nu,n})] \delta(r - r_{\nu,n}), \quad (3.7)$$

en donde (r, θ) son las coordenadas polares, $\beta_{\nu,n}$ es el ancho del anillo, $\phi_{\nu,n}$ es una fase constante que puede tomar los valores 0 o π/ν y $r_{\nu,n} = \alpha_{\nu,n} R/\alpha_{\nu,N}$, en donde R es el radio de la zona activa del LC-SLM. La señal óptica en el plano focal de la lente se da a través de la transformada de Fourier

Generación de un vórtice óptico usando técnicas de recuperación de fase
3.2 Esquemas de generación de vórtices ópticos: coherentes y parcialmente coherentes

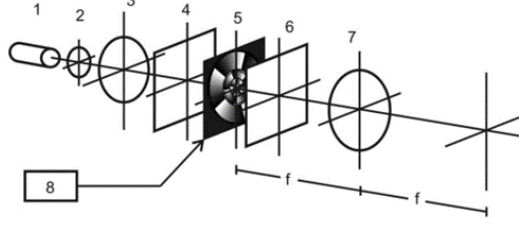


Figura 3.2: Montaje experimental para generar un vórtice óptico perfecto. 1 Láser, 2 y 3 forman un expansor de haz, 4 y 6 son polarizadores, 5 LC-SLM, 7 es una lente esférica con distancia focal f y 8 es una computadora.

$$U_\nu(\rho, \theta) = \int_0^\infty \int_0^{2\pi} t_\nu(r, \varphi) \exp\left(-i\frac{2\pi}{\lambda f} r \rho \cos(\varphi - \theta)\right) r dr d\varphi, \quad (3.8)$$

donde λ es la longitud de onda de iluminación. Sustituyendo la transmitancia de la Ec. (3.7) en la Ec. (3.8), obtenemos

$$U_\nu(\rho, \theta) = \exp(-i\nu\theta) \sum_{n=1}^N \alpha_{\nu,n}^2 \beta_{\nu,n} \exp(-i\nu\varphi_{\nu,n}) J_\nu\left(\frac{2\pi R \alpha_{\nu,n} \rho}{\lambda f \alpha_{\nu,N}}\right), \quad (3.9)$$

en donde se ha utilizado la propiedad de filtración de la función δ y la identidad de Bessel

$$J_\nu(x) = \frac{i^{-\nu}}{2\pi} \int_0^{2\pi} \exp(i\nu\theta) \exp(ix \cos\theta) d\theta. \quad (3.10)$$

Ahora, haciendo

$$\alpha_\nu = \frac{\lambda f \alpha_{\nu,N}}{2\pi R}, \quad (3.11)$$

y aplicando

$$\beta_{\nu,n} = \frac{|J_\nu(\alpha_{\nu,n} \rho_0 / a_\nu)|}{\alpha_{\nu,n}^2 [J_{\nu+1}(\alpha_{\nu,n})]^2}, \quad (3.12)$$

$$\varphi_{\nu,n} = \begin{cases} 0 & \text{para } J_\nu(\alpha_{\nu,n} \rho_0 / a_\nu) \geq 0 \\ \pi/\nu & \text{para } J_\nu(\alpha_{\nu,n} \rho_0 / a_\nu) < 0, \end{cases} \quad (3.13)$$

se encuentra finalmente

$$U_\nu(\rho, \theta) \propto \exp(i\nu\theta) \sum_{n=1}^N \frac{J_\nu(\alpha_{\nu,n} \rho_0 / a_\nu)}{[J_{\nu+1}(\alpha_{\nu,n})]^2} J_\nu\left(\alpha_{\nu,n} \frac{\rho}{a_\nu}\right). \quad (3.14)$$

La Ec. (3.14) representa la aproximación finita de la representación del vórtice perfecto de la expresión extendida en el rango entero $(0, \infty)$ de la coordenada ρ .

Los vórtices generados utilizando el método propuesto son ilustrados en la Fig. (3.3). De las imágenes se puede observar que los vórtices generados presentan irregularidades de intensidad a lo largo de la dirección azimutal. Se puede observar también que presentan un centro oscuro en el centro correspondiente a la región de la singularidad de fase. En todos los casos se aprecia

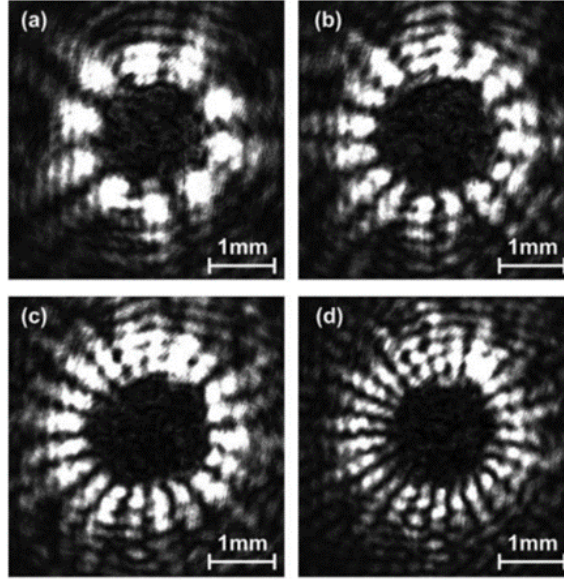


Figura 3.3: Vórtices generados experimentalmente para cargas topológicas de a) $\nu = 5$, b) $\nu = 8$, c) $\nu = 10$ y d) $\nu = 10$.

que dicha región es de dimensiones similares lo cual es una característica predicha para un vórtice perfecto. Sin embargo, aún cuando la simulación computacional confirma la teoría, en el caso experimental los vórtices generados distan mucho de corresponder al caso teórico. En parte esto se debe a la discretización de la señal de fase y la capacidad del modulados para desplegar dicha fase. Por tanto es posible observar que aún cuando se conoce la fase necesaria para generarlos, en la práctica no es una tarea fácil. En la siguiente sección describiremos una técnica que reproduce mejor la generación experimental del vórtice perfecto.

Técnica alternativa para la generación de un vórtice óptico perfecto

Con el fin de mejorar la calidad del vórtice generado con la técnica anterior, se propuso una nueva técnica cuya importancia radica en la simplicidad de su realización práctica y la alta calidad de los resultados obtenidos. Para obtener un vórtice perfecto, es posible calcular la transformada de Fourier inversa de la Ec. (3.3) para obtener que la señal óptica de entrada del sistema esta dada por [31]

$$U(r, \varphi) \propto J_\nu \left(\frac{2\pi}{\lambda f} r \rho_0 \right) \exp(i\nu\varphi). \quad (3.15)$$

Luego, admitiendo la aproximación de la función de Bessel de la ecuación anterior para valores positivos de ν descrita por:

$$J_\nu \left(\frac{2\pi}{\lambda f} r \rho_0 \right) \approx \sum_n (-1)^{n-1} \text{rect} \left(\frac{r - r_{\nu,n}}{\Delta r_{\nu,n}} \right), \quad (3.16)$$

en donde $\text{rect}(\cdot)$ es la función rectángulo, $r_{\nu,n}$ es la n -ésima raíz de la ecuación

$$\frac{d}{dr} J_\nu \left(\frac{2\pi}{\lambda f} r \rho_0 \right) = 0, \quad (3.17)$$

y

Generación de un vórtice óptico usando técnicas de recuperación de fase
 3.2 Esquemas de generación de vórtices ópticos: coherentes y parcialmente coherentes

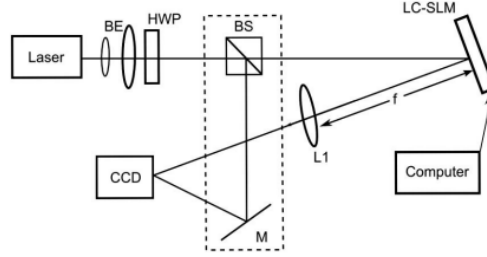


Figura 3.4: Montaje experimental para generar los vórtices ópticos. BE, expansor de haz; HWP; placa de media onda; BS, divisor de haz; y L y M son lentes y espejos respectivamente.

$$\Delta r_{\nu,n} \propto \frac{|J_{\nu}(2\pi r_{\nu,n}\rho_0/\lambda f)|}{J_{\nu}(2\pi r_{\nu,n}\rho_0/\lambda f)}. \quad (3.18)$$

Así, el campo de la Ec. (3.15) puede ser reproducido aproximadamente transmitiendo una onda plana a través de una pantalla de fase formada por un conjunto de anillos concéntricos con transmitancia de amplitud dada por:

$$T(r, \varphi) = \exp\left(i\frac{2\pi}{\lambda f}x'_0x\right) \sum_n \text{rect}\left(\frac{r - r_{\nu,n}}{\Delta r_{\nu,n}}\right) \exp[i\nu(\varphi - \varphi_{\nu,n})], \quad (3.19)$$

separados por un conjunto de anillos concéntricos complementarios con transmitancia de amplitud

$$T'(r, \varphi) = \exp\left(-i\frac{2\pi}{\lambda f}x'_0x\right) \left[1 - \sum_n \text{rect}\left(\frac{r - r_{\nu,n}}{\Delta r_{\nu,n}}\right)\right], \quad (3.20)$$

En la Ec. (3.19) $\varphi_{\nu,n}$ toma los valores 0 para n pares y ϕ/ν para n impares. Se encuentra entonces que la señal centrada en el punto ($x' = -x'_0, y' = 0$) en la salida del sistema óptico es

$$U(\rho, \theta) = \exp(i\nu\theta) \sum_n (-1)^{n-1} \int_{r_{\nu,n}-\Delta r_{\nu,n}/2}^{r_{\nu,n}+\Delta r_{\nu,n}/2} J_{\nu}\left(\frac{2\pi}{\lambda f}r\rho\right) r dr, \quad (3.21)$$

cuya distribución de intensidad es

$$I(\rho) = \left|\sum_n (-1)^{n-1} \int_{r_{\nu,n}-\Delta r_{\nu,n}/2}^{r_{\nu,n}+\Delta r_{\nu,n}/2} J_{\nu}\left(\frac{2\pi}{\lambda f}r\rho\right) r dr\right|^2, \quad (3.22)$$

con N el parámetro de truncamiento de la sumatoria.

Los resultados de esta técnica fueron mostrados usando el arreglo de la figura Fig(3.4). En esta técnica se emplea nuevamente un sistema de transformada de Fourier solamente que en este caso el objeto corresponde a la onda que emerge del modulador de reflexión donde se han codificado las transmitancias dadas por las Ecs. (3.19-3.20). En la figura se observa un divisor de haz empleado para lograr la interferencia del campo proveniente del modulador con una onda plana a fin de mostrar los típicos interferogramas con bifurcaciones que revelan la presencia del vórtice con carga ν .

3.2.1. Vórtices parcialmente coherentes

En el ejemplo anterior se mostró que para generar un vórtice óptico perfecto se requiere codificar una fase espacialmente definida en secciones anulares que siguen la aproximación rectangular de la función de Bessel.

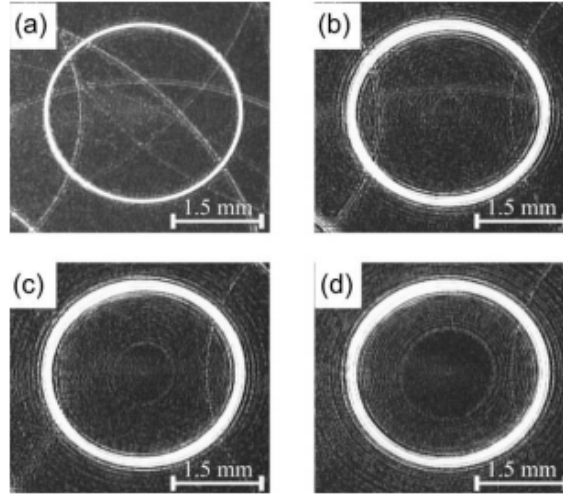


Figura 3.5: Vórtices generados experimentalmente con un radio de $\rho_0 = 1,5mm$ y cargas topológicas de a) $\nu = 1$, b) $\nu = 20$, c) $\nu = 40$ y d) $\nu = 6$

En esta sección presentaremos los fundamentos de una clase completamente diferente de vórtices conocidos como parcialmente coherentes. A diferencia de los anteriores que se generan a partir de funciones completamente deterministas. Los parcialmente coherentes se generan a partir de un ensamble estadístico que determina las propiedades del campo. Dichos haces pueden representarse como una superposición incoherente de modos Bessel completamente coherentes [33], todos con el mismo frente de onda helicoidal. Una característica importante de éste tipo de vórtices es que son adifraccionales [32]. Para entender la teoría involucrada en la generación de éste tipo de haces, es necesario introducir algunos conceptos [34].

Un campo óptico transversal con coherencia parcial de segundo orden es descrito por una función de correlación conocida como la densidad espectral $W(\rho_1, \rho_2) = \langle E^*(\rho_1)E(\rho_2) \rangle$ con $E(\rho, \nu)$ un campo eléctrico transversal, en adelante omitiremos por simplicidad la dependencia explícita en la frecuencia ν . Esta expresión contiene toda la información necesaria para describir el comportamiento del campo. Bajo ciertas condiciones dicha densidad espectral puede expresarse mediante una expansión en serie conocida como serie de Mercer, la cual está dada por la expresión

$$W(\rho_1, \rho_2) = \sum_n \lambda_n \varphi_n^*(\rho_1) \varphi_n(\rho_2), \quad (3.23)$$

donde λ_n son los eigenvalores y $\varphi_n(\rho)$ son las eigenfunciones de la ecuación integral homogénea de Fredholm de segundo tipo

$$\int_D W(\rho_1, \rho_2) \varphi_n(\rho_1) d\rho_1 = \lambda_n \varphi_n(\rho_2), \quad (3.24)$$

con D el dominio en el cual está definida la fuente y siendo λ_n necesariamente no negativos. Además $\varphi_n(\rho)$ forma un conjunto ortonormal en el dominio de la fuente. Cada término de la suma de la Ec. (3.23) está asociado con un modo completamente coherente $\varphi_n(\rho)$ del campo, por esta razón la Ec. (3.23) es comúnmente llamada *representación de modos coherentes* de un campo parcialmente coherente. Ahora, consideramos un proceso aleatorio formado por un ensamble estadístico de realizaciones, en donde cada realización es representada por:

$$U(\rho) = \sum_n C_n \varphi_n(\rho). \quad (3.25)$$

Dicho campo tiene una densidad espectral dada por

Generación de un vórtice óptico usando técnicas de recuperación de fase
 3.2 Esquemas de generación de vórtices ópticos: coherentes y parcialmente coherentes

$$W(\rho_1, \rho_2) = \langle U^*(\rho_1)U(\rho_2) \rangle = \sum_n \sum_m \langle C_n^* C_m \rangle \varphi_n^*(\rho_1) \varphi_m(\rho_2), \quad (3.26)$$

en donde $\langle \cdot \rangle$ denota un promedio estadístico en el ensamble de realizaciones aleatorias. Los modos $\varphi_n(\rho)$ son elegidos de la forma

$$\varphi_n(\rho) = \frac{J_\nu(\alpha_{\nu,n} \frac{\rho}{a})}{\sqrt{\pi a} J_{\nu+1}(\alpha_{\nu,n})} \exp(i\nu\theta) \quad 0 \leq \rho \leq a, \quad (3.27)$$

los cuales forman un conjunto ortonormal de funciones de Bessel. Las variables aleatorias C_n satisfacen la condición

$$\langle C_n^* C_m \rangle = \lambda_n \delta_{n,m}, \quad (3.28)$$

en donde $\delta_{n,m}$ es una delta de Kronecker. A partir de éstas condiciones, substituyendo las Ecs. (3.27-3.28) en la Ec. (3.26) se obtiene

$$W(\rho_1, \rho_2) = \sum_n \lambda_n \frac{J_\nu(\alpha_{\nu,n} \frac{\rho_1}{a}) J_\nu(\alpha_{\nu,n} \frac{\rho_2}{a})}{\pi a^2 J_{\nu+1}^2(\alpha_{\nu,n})} \times \exp[i\nu(\theta_2 - \theta_1)], \quad (3.29)$$

y la densidad espectral de potencia, o intensidad óptica está dada por

$$S(\rho) = \sum_n \lambda_n \frac{J_\nu^2(\alpha_{\nu,n} \frac{\rho}{a})}{\pi a^2 J_{\nu+1}^2(\alpha_{\nu,n})}. \quad (3.30)$$

La Ec.(3.29), con los valores arbitrarios de λ_n puede identificarse como un haz con vórtice parcialmente coherente en modos de Bessel. Además puede demostrarse que la familia de haces definidos por la Ec. (3.29) se propagan sin difracción. En la práctica, este tipo de vórtices pueden ser generados aproximadamente en un sistema transformador de Fourier por medio de un SLM. El esquema para formar este tipo de vórtices se muestra en la Fig. (3.6). En la Fig. 3.7 se observan vórtices generados por esta técnica para algunos valores de la carga topológica.

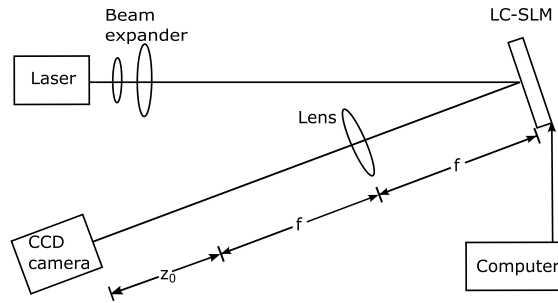


Figura 3.6: Montaje experimental para la generación de vórtices con estructura modo Bessel

Como hemos visto, este tipo de campos se originan a partir de un ensamble estadístico compuesto por un gran número de realizaciones. Teóricamente el ensamble consiste de todas las posibles manifestaciones que puede tener el campo óptico. Esto es, cada posible configuración del campo está representadas por una realización. Debido a esto, no es posible hablar de una sola fase asociada al vórtice óptico sino es posible que existan una infinidad de fases. Sin embargo para poder generar los vórtices mediante un sistema de transformada de Fourier, se emplea un detector CCD con resolución espacial que genera un promedio de las realizaciones del campo. Dicho promedio es proporcional a la intensidad óptica. Por lo anterior es posible observar que este tipo de vórtices son

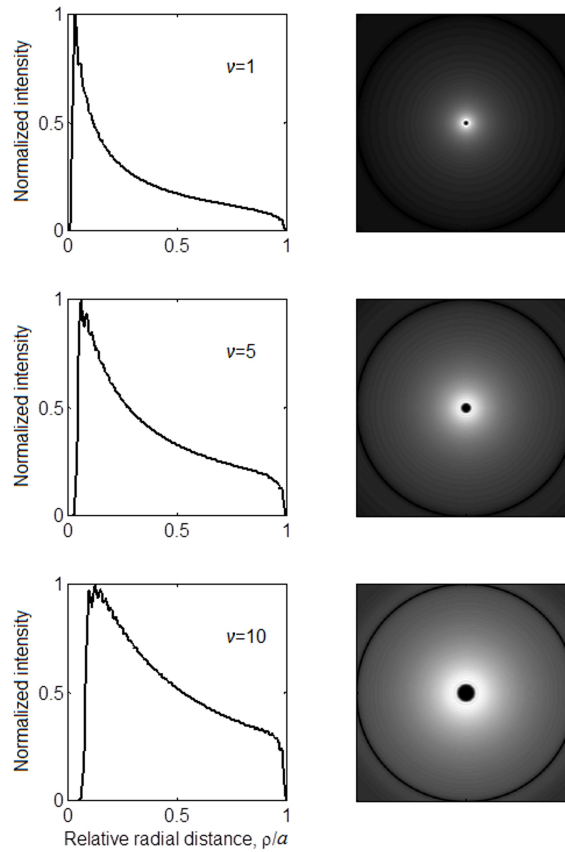


Figura 3.7: Resultados experimentales

de interés en diversas aplicaciones sin embargo su generación no es una tarea fácil. En la siguiente sección emplearemos un método iterativo para recuperar una fase que permita generar este tipo de vórtices sin necesidad de hacer un promedio estadístico, a partir del conocimiento o medición del módulo del campo deseado.

3.3. Síntesis iterativa de un vórtice óptico

Las técnicas de generación de vórtices ópticos vistas anteriormente requieren de la codificación de la fase de la señal, sin embargo, como hemos visto, en algunas ocasiones éste proceso puede resultar complicado. Por otra parte, los algoritmos iterativos tipo Gerchberg-Saxton, resultan ser una práctica herramienta para obtener una fase que permita reconstruir, en este caso, el vórtice óptico deseado.

En éste trabajo se propone recuperar la fase tanto de un vórtice óptico perfecto como de un vórtice óptico parcialmente coherente implementando el algoritmo de Gerchberg-Saxton (visto en el capítulo anterior) en Matlab. La información que se supone conocida es la intensidad del campo en el plano del objeto y la intensidad del campo en el plano de Fourier, dicha información representa las restricciones impuestas en el plano del objeto y en el plano de difracción respectivamente. La transformada directa e inversa de Fourier son aplicadas en ambos planos con cada iteración del algoritmo. Al satisfacerse la condición establecida para la detención del algoritmo (en nuestro caso un determinado número de iteraciones), la distribución de fase obtenida es utilizada posteriormente para la reconstrucción del vórtice óptico.

Si suponemos un haz monocromático, coherente y que puede ser descrito mediante una función compleja, según la teoría escalar de difracción vista en el Capítulo 1, el campo en el plano del objeto puede ser escrito como:

$$U(u, v) = A(u, v) \exp[i\phi(u, v)], \quad (3.31)$$

en donde (u, v) son las coordenadas en el plano del objeto, $A(u, v)$ es la amplitud del campo y ϕ representa la fase. De manera similar, el campo en el plano de difracción, puede escribirse como

$$V(x, y) = B(x, y) \exp[i\psi(x, y)], \quad (3.32)$$

en donde (x, y) son las coordenadas en el plano de difracción, $B(x, y)$ la amplitud y $\psi(x, y)$ la fase. El procedimiento para el cálculo de la fase del vórtice óptico puede ser listado como sigue:

- 1.- Se eligen la amplitud $A_0(u, v)$ y fase inicial $\phi_0(u, v)$ como entrada al algoritmo

$$U_0 = A_0(u, v) \exp[i\phi_0(u, v)]. \quad (3.33)$$

- 2.- Calculamos la FFT de la función resultante.

3.- La amplitud compleja resultante en el plano de difracción es reemplazada por la restricción impuesta en el mismo, es decir:

$$V_1 = B_0(x, y) \exp[i\psi(x, y)] \rightarrow V_1 = B(x, y) \exp[i\psi(x, y)], \quad (3.34)$$

en donde V_1 es el resultado de aplicar la FFT a la función U_0 y $B(x, y)$ es la restricción impuesta en el plano de difracción.

- 4.- Calculamos la IFFT de la Ec. (3.34)

$$U_1 = IFFT[V_1]. \quad (3.35)$$

5.- La nueva amplitud resultante en el plano del objeto es reemplazada por la restricción impuesta en dicho plano, se tiene que

$$U_1 = A_1(u, v) \exp[i\phi_1(u, v)] \rightarrow U_1 = A(u, v) \exp[i\phi_1(u, v)]. \quad (3.36)$$

6.- El proceso se repite del paso 2 al paso 5 hasta cumplir la condición impuesta para la detención del algoritmo.

Es importante considerar que la fase obtenida a través de la implementación del algoritmo, está envuelta, es decir, se encuentra empaquetada entre los valores $-\pi$ y π . Este envolvimiento ocurre al utilizar la función *angle* en Matlab para la obtención de la fase buscada, y que a su vez hace uso de la función tangente, con una periodicidad de π radianes. sin embargo esto no es un inconveniente para la realización del algoritmo ya que los valores de fase no cambian con la función *angle*, sino que solamente se mantienen en un intervalo fijo en donde la función tangente es biyectiva. De esta manera, para estudiar la fase obtenida al finalizar el algoritmo, es necesario el uso de algún procedimiento para su desenvolvimiento. Sin embargo, en nuestro en el siguiente capítulo mostramos como es posible implementar el algoritmo en la práctica para dos tipos de vórtices completamente diferentes y su capacidad para obtener la fase.

3.3.1. Análisis de error

Para monitorear el comportamiento del algoritmo en función del acercamiento a la intensidad deseada es posible calcular el error cuadrático medio (ECM) en cada reconstrucción. Dicho error nos permite estimar un valor positivo específico de la diferencia entre la intensidad estimada del vórtice y la intensidad medida. El ECM está definido por

$$ECM = \frac{1}{n} \sum_{i=1}^n (I_i - \hat{I}_i)^2. \quad (3.37)$$

donde I_i es la intensidad medida en cada entrada de la matriz de datos e \hat{I}_i es la intensidad estimada en cada entrada correspondiente de la matriz asociada a la intensidad reconstruida. En el siguiente capítulo estimaremos la fase asociada a la generación de dos tipos de vórtices y en cada caso estimaremos el ECM para cada iteración.

Capítulo 4

Resultados numéricos

Con el fin de comprobar el comportamiento del algoritmo y obtener la fase asociada a una intensidad medida elegimos dos tipos de vórtices, uno completamente coherente dado por el vórtice perfecto y otro parcialmente coherente descritos en el capítulo 3. En ambos casos la distribución de intensidad del vórtice es obtenida en el plano de Fourier. En el primer caso, la intensidad óptica corresponde al cuadrado del módulo de la transformada de Fourier de la función de Bessel de orden ν multiplicada por un termino de fase definido por la ecuación 3.5. En el segundo empleamos la intensidad óptica del vórtice generado en [33] para una carga topológica de $\nu = 10$?

4.1. Vórtice óptico perfecto

En el caso coherente suponemos que el termino de fase que genera el vórtice es de la forma dada por la ecuación 3.5 pero no conocemos el valor de la carga topológica y por tanto se propone $\nu = 3$ como carga de prueba. Al calcular el módulo de la función compleja generada obtenemos la restricción que será impuesta en el plano del objeto $f = |g_\nu(\rho, \theta)|$. Por otro lado, al calcular el módulo de la FFT de la función compleja mencionada anteriormente, obtenemos la restricción impuesta en el plano de Fourier $F = |FFT[g_\nu(\rho, \theta)]|$. La intensidad y fase del vórtice óptico propuesto (es decir, el vórtice teórico) se ilustran en la Fig.(4.1).

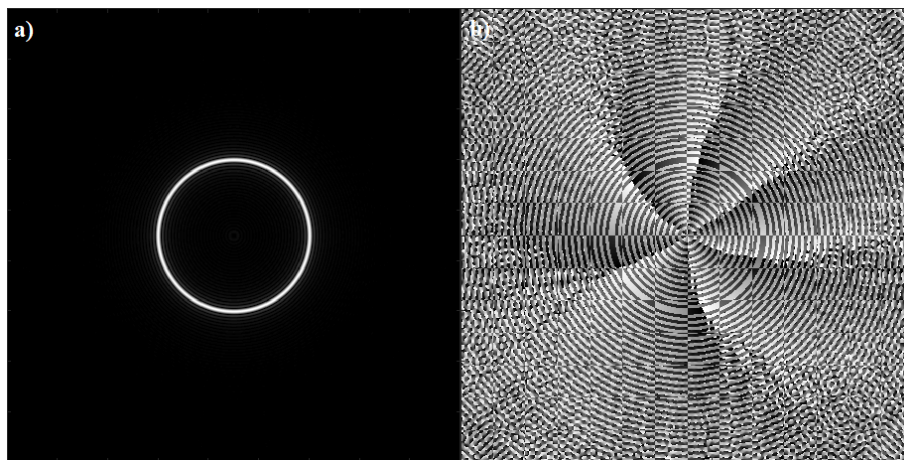


Figura 4.1: a) Intensidad y b) fase del vórtice óptico perfecto teórico, con carga topológica $\nu = 3$.

A fin de comparar la convergencia del algoritmo y el número de iteraciones utilizadas, se proponen

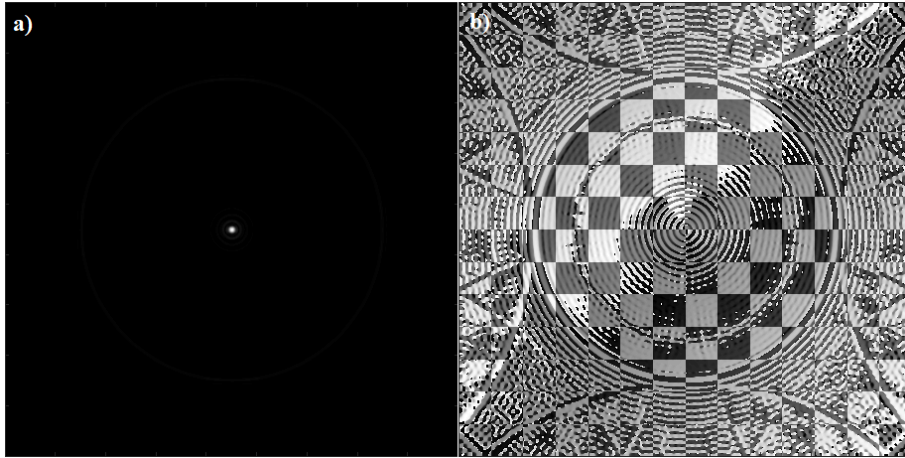


Figura 4.2: Recuperación de a) amplitud y b) fase de un VOP para 1 iteración.

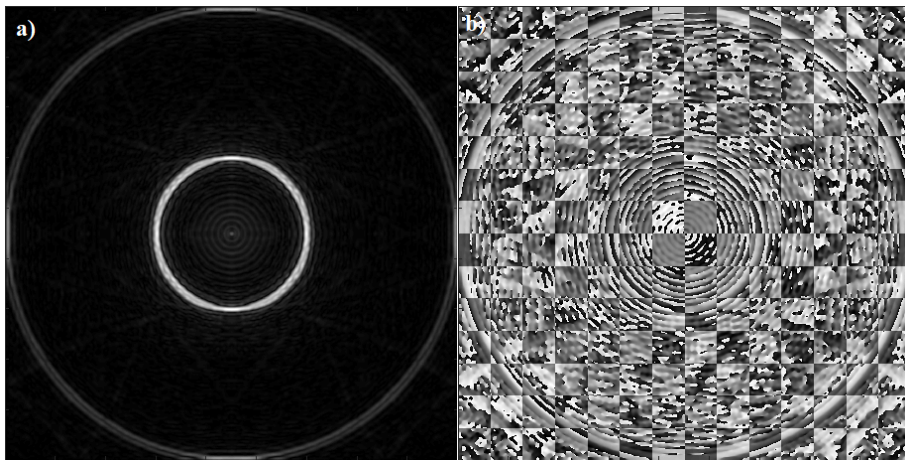


Figura 4.3: Recuperación de a) amplitud y b) fase de un VOP para 5 iteraciones.

dos fases. Una de ellas es una función cuadrática mientras que la otra es una fase con un perfil helicoidal. Se implementa el algoritmo mediante un ciclo *while* para distintas iteraciones y ambas aproximaciones iniciales de las fases mencionadas.

En el caso de un vórtice óptico perfecto con fase inicial dada por la función cuadrática x^2 , los resultados obtenidos se muestran en las Fig. (4.2)-(4.11). Para el caso en donde la fase inicial está dada por $\arg(Z)$ (en donde $Z = x + iy$), los resultados obtenidos se muestran en las Fig. (4.12)-(4.24).

Como se puede observar a partir de los resultados obtenidos, al utilizar una fase cuadrática como primera aproximación el algoritmo converge de más rápida dentro de las primeras 60 iteraciones, la reconstrucción de vórtice se alcanza en las primeras 100 iteraciones y a partir de aquí los cambios son mínimos. A diferencia del caso anterior, para una fase con un perfil helicoidal como primera aproximación la convergencia del algoritmo resulta ser un poco más lenta y la reconstrucción del vórtice se alcanza a partir de 150 iteraciones. A pesar de que la convergencia del algoritmo es distinta para cada una de las fases propuestas, es posible obtener la fase que permite reconstruir el vórtice óptico perfecto en ambos casos.

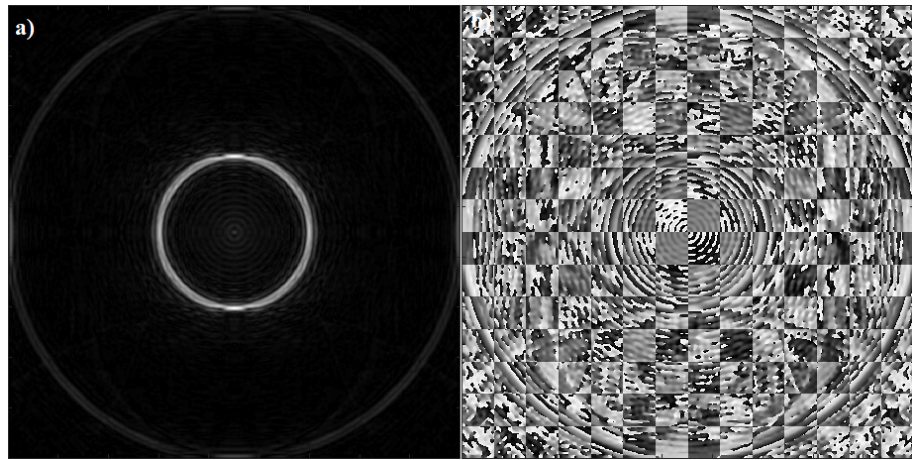


Figura 4.4: Recuperación de a) amplitud y b) fase de un VOP para 10 iteraciones.

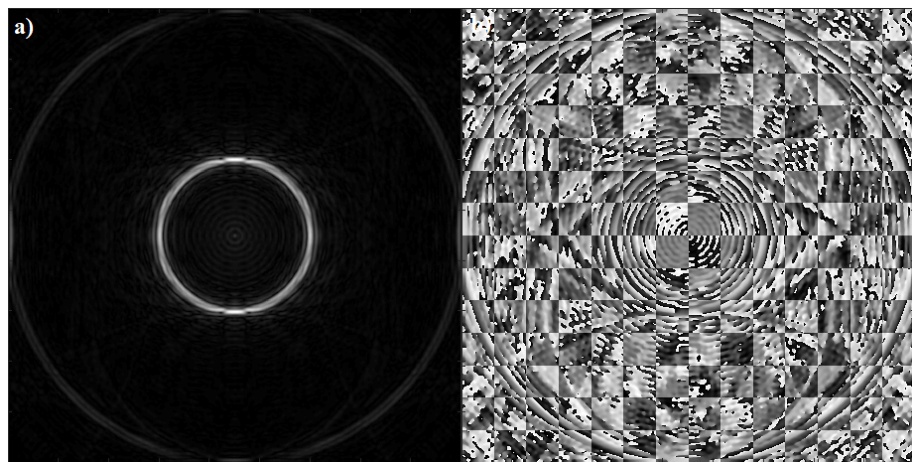


Figura 4.5: Recuperación de a) amplitud y b) fase de un VOP para 20 iteraciones.

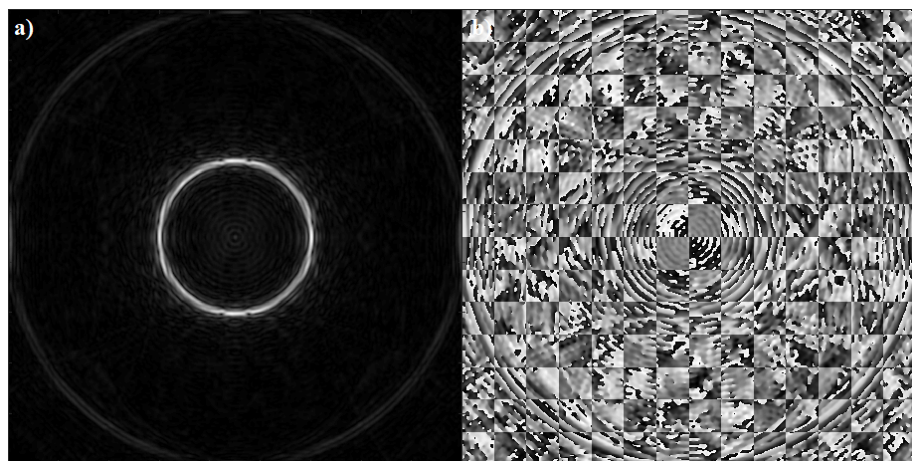


Figura 4.6: Recuperación de a) amplitud y b) fase de un VOP para 30 iteraciones.

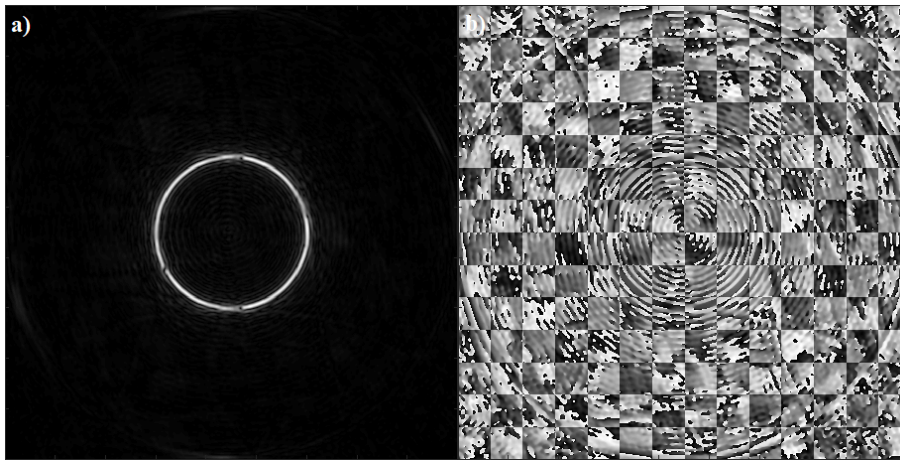


Figura 4.7: Recuperación de a) amplitud y b) fase de un VOP para 50 iteraciones.

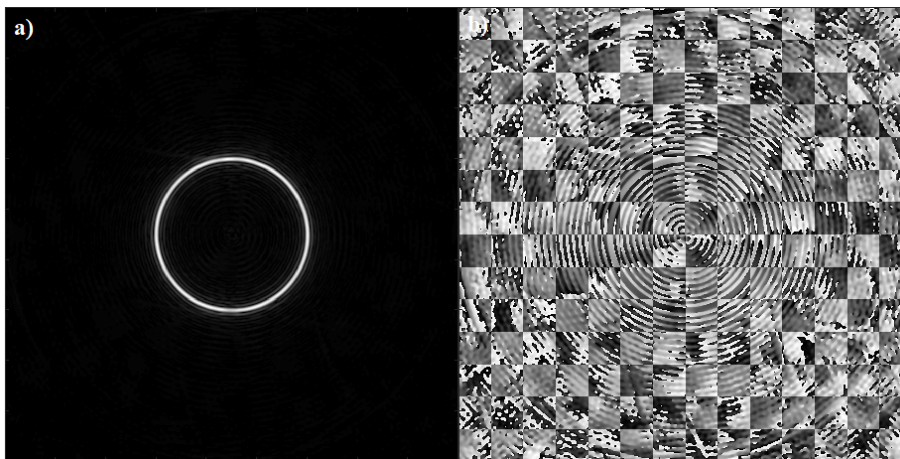


Figura 4.8: Recuperación de a) amplitud y b) fase de un VOP para 60 iteraciones.

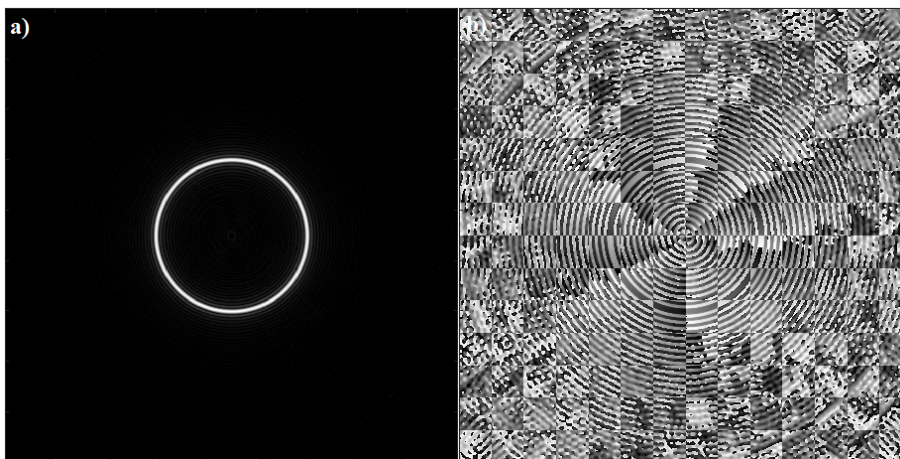


Figura 4.9: Recuperación de a) amplitud y b) fase de un VOP para 80 iteraciones.

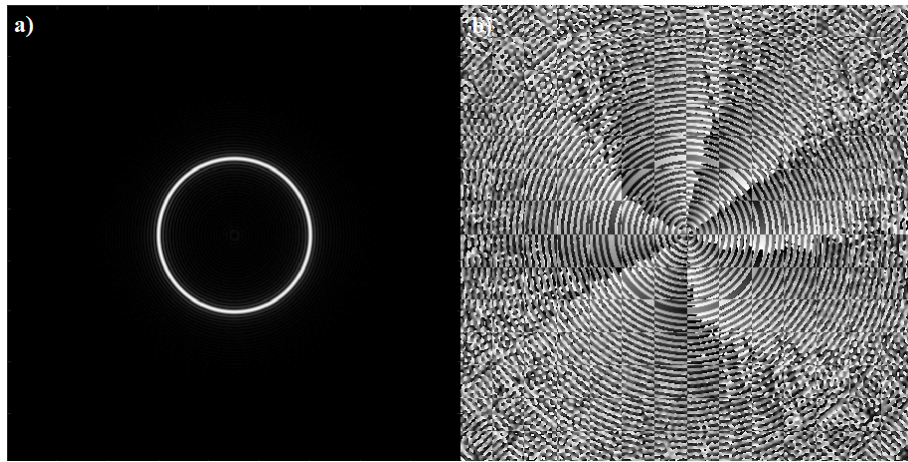


Figura 4.10: Recuperación de a) amplitud y b) fase de un VOP para 100 iteraciones.

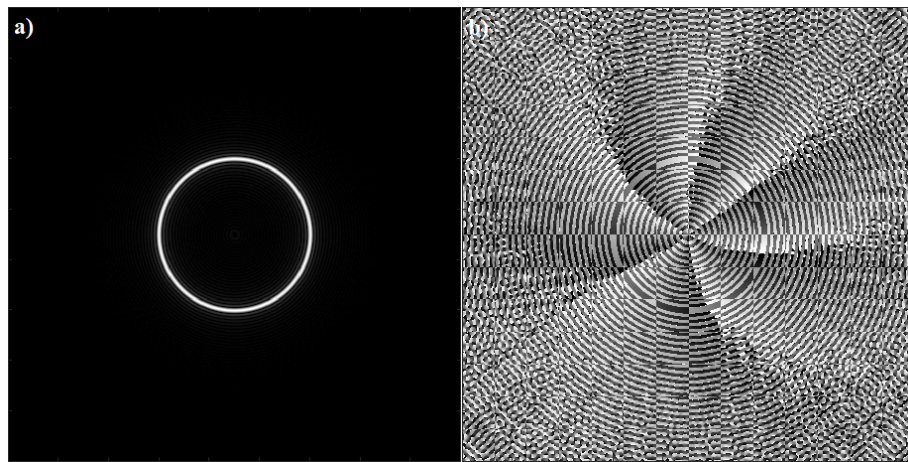


Figura 4.11: Recuperación de a) amplitud y b) fase de un VOP para 200 iteraciones.

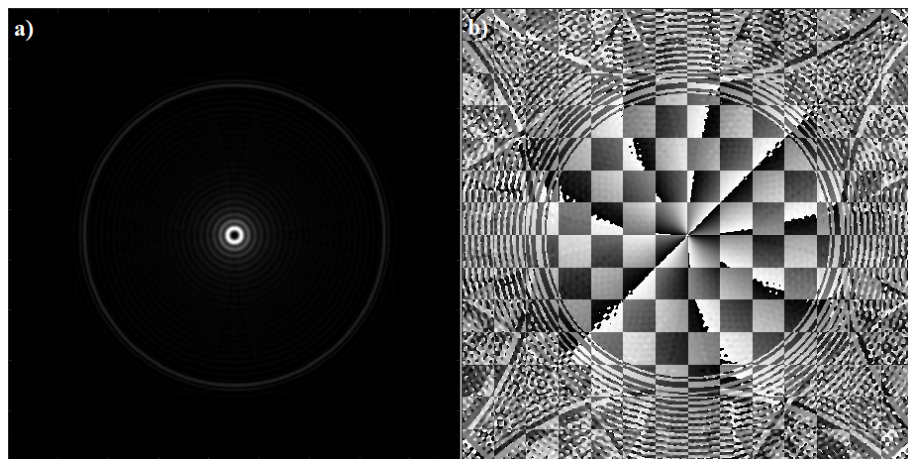


Figura 4.12: Recuperación de a) amplitud y b) fase de un VOP para 1 iteración.

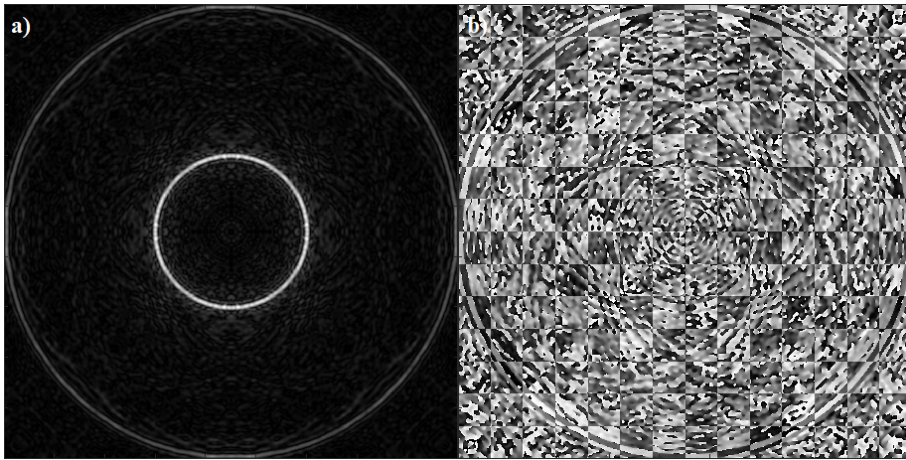


Figura 4.13: Recuperación de a) amplitud y b) fase de un VOP para 5 iteraciones.

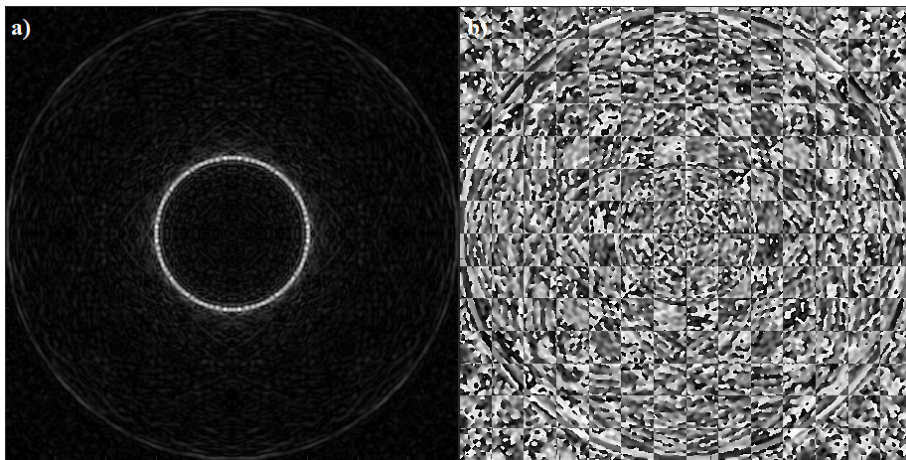


Figura 4.14: Recuperación de a) amplitud y b) fase de un VOP para 10 iteraciones.

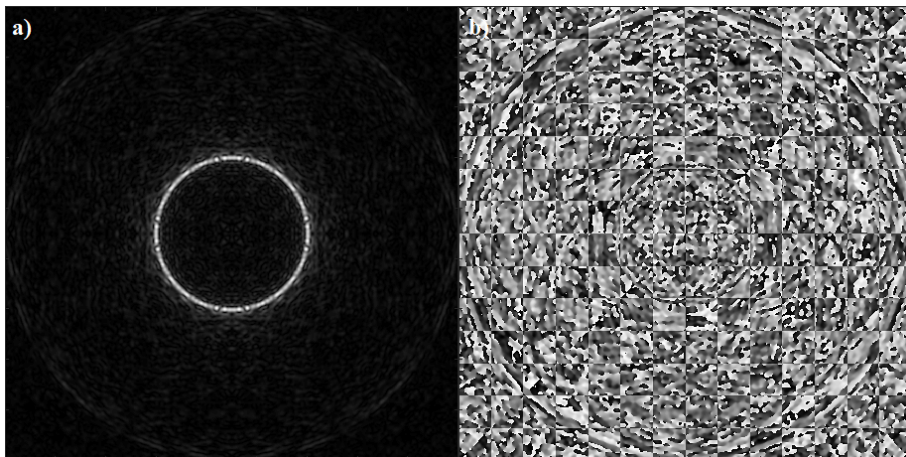


Figura 4.15: Recuperación de a) amplitud y b) fase de un VOP para 20 iteraciones.

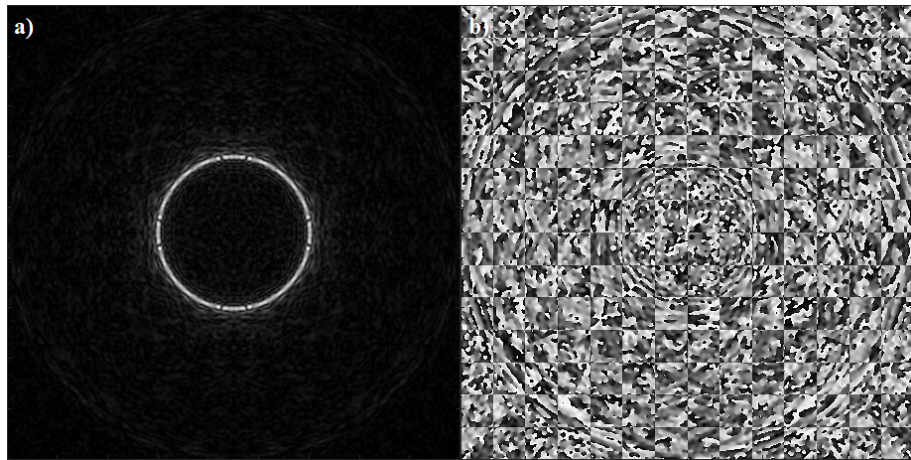


Figura 4.16: Recuperación de a) amplitud y b) fase de un VOP para 30 iteraciones.

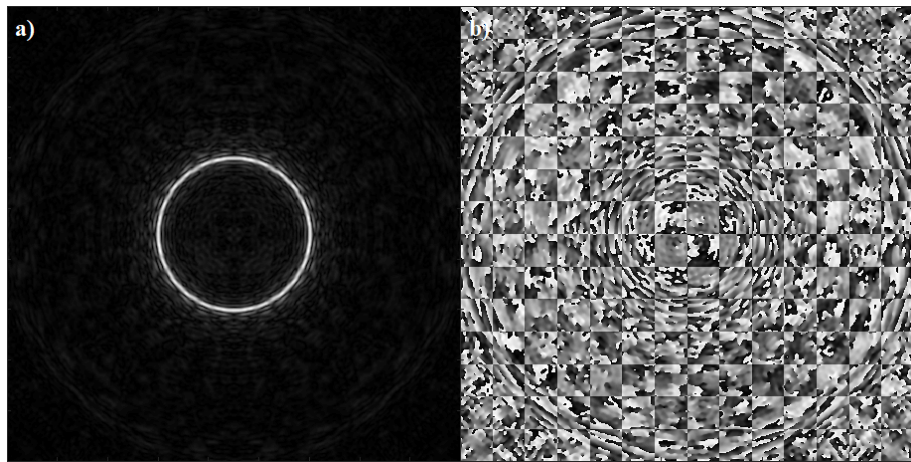


Figura 4.17: Recuperación de a) amplitud y b) fase de un VOP para 50 iteraciones.

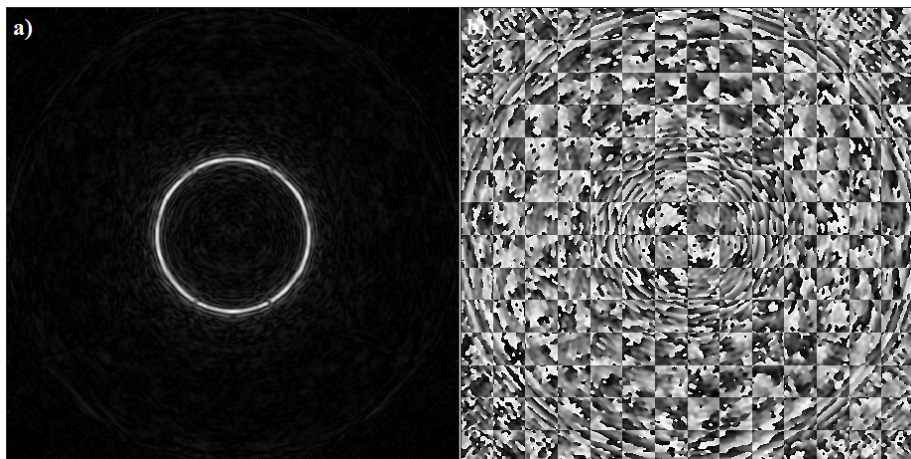


Figura 4.18: Recuperación de a) amplitud y b) fase de un VOP para 100 iteraciones.

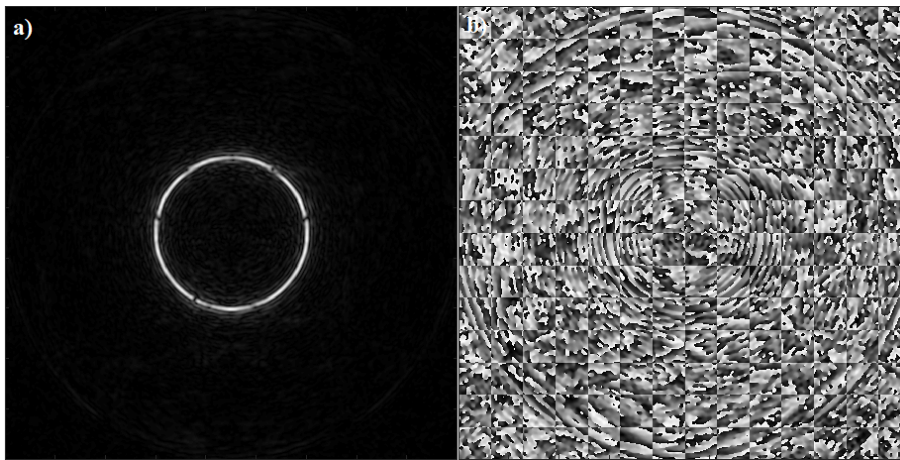


Figura 4.19: Recuperación de a) amplitud y b) fase de un VOP para 125 iteraciones.

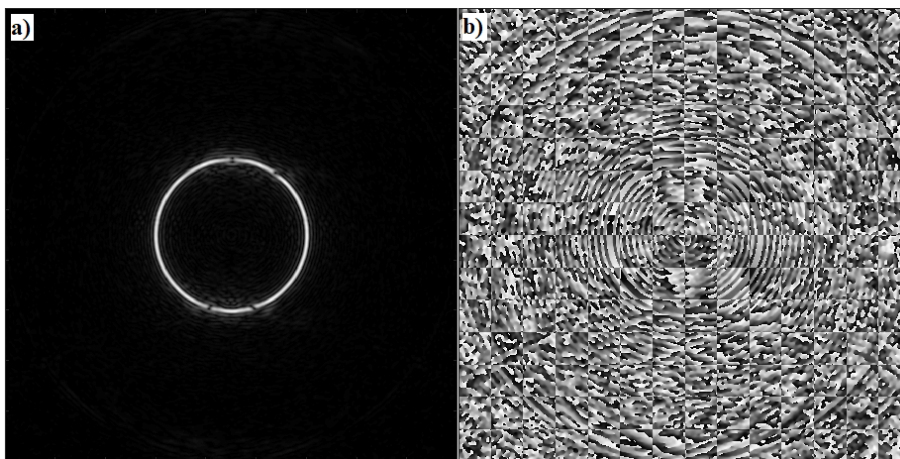


Figura 4.20: Recuperación de a) amplitud y b) fase de un VOP para 135 iteraciones.

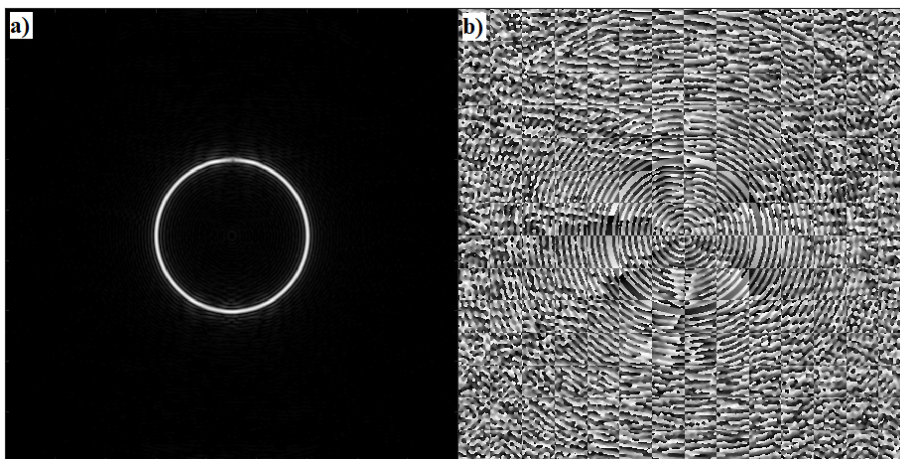


Figura 4.21: Recuperación de a) amplitud y b) fase de un VOP para 145 iteraciones.

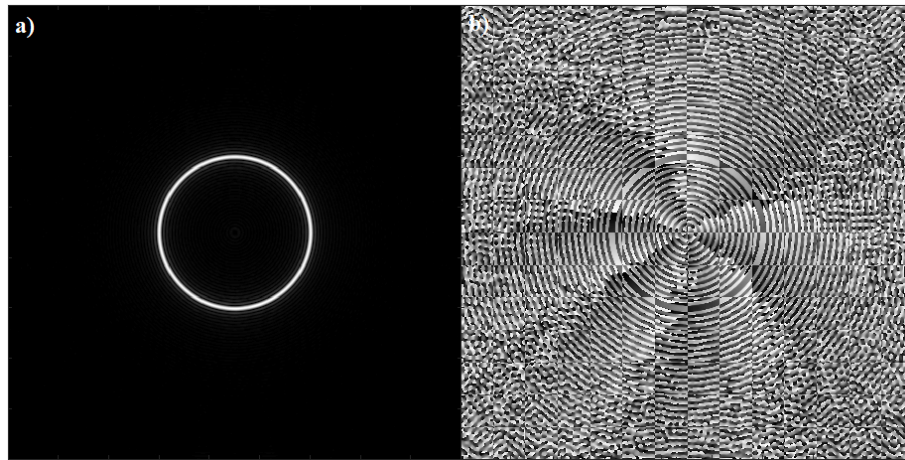


Figura 4.22: Recuperación de a) amplitud y b) fase de un VOP para 150 iteraciones.

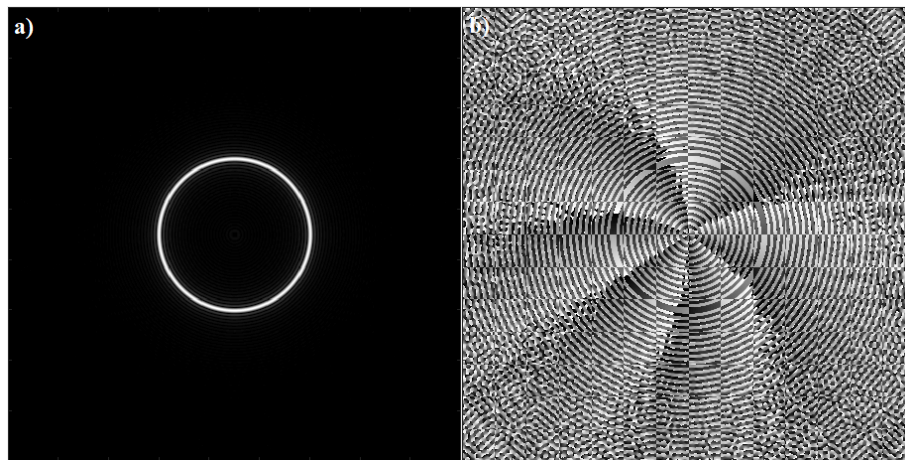


Figura 4.23: Recuperación de a) amplitud y b) fase de un VOP para 175 iteraciones.

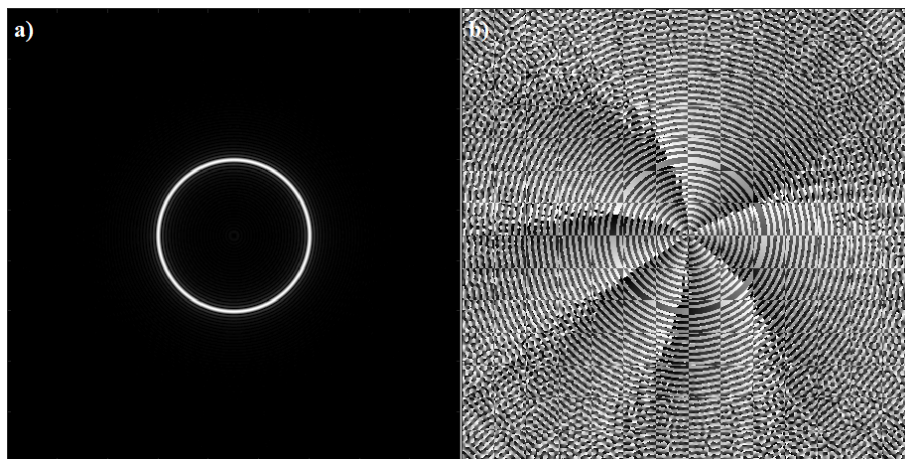


Figura 4.24: Recuperación de a) amplitud y b) fase de un VOP para 200 iteraciones.

4.2. Vórtice óptico parcialmente coherente

Para el caso de un vórtice parcialmente coherente, la restricción impuesta en el plano del objeto es una función definida por un término de fase lineal x definido en un dominio circular de tal forma que la función es cero fuera del dominio y posee valores de la fase inicial dentro del dominio. En el plano de difracción, la restricción es la FFT de dicha intensidad, es decir la intensidad medida del vórtice. Los resultados obtenidos se muestran en las Fig. 4.26-4.38. Como se observa, en este caso el vórtice óptico es reconstruido después de 20 iteraciones y a partir de aquí los cambios son mínimos. Tal como indican los resultados, el algoritmo permite encontrar una fase con la cual es posible reconstruir el vórtice parcialmente coherente.

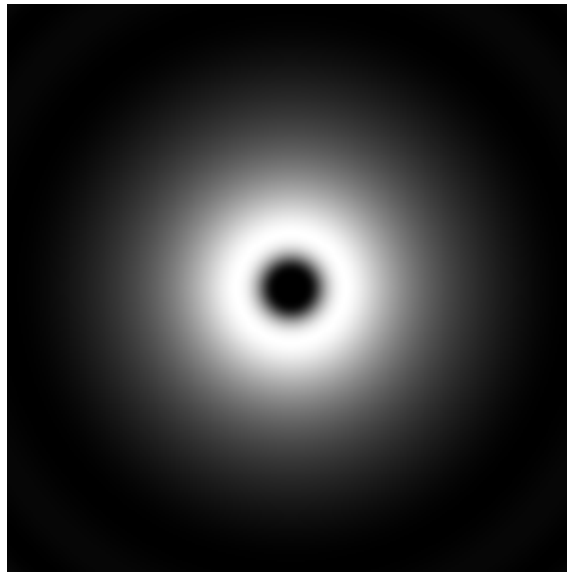


Figura 4.25: Distribución de intensidad de un VPC

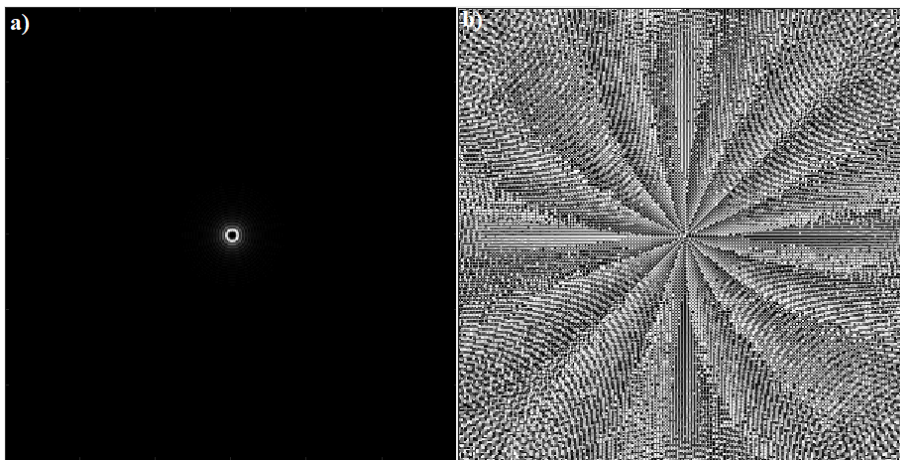


Figura 4.26: Recuperación de a) amplitud y b) fase de un VPC para 1 iteración.

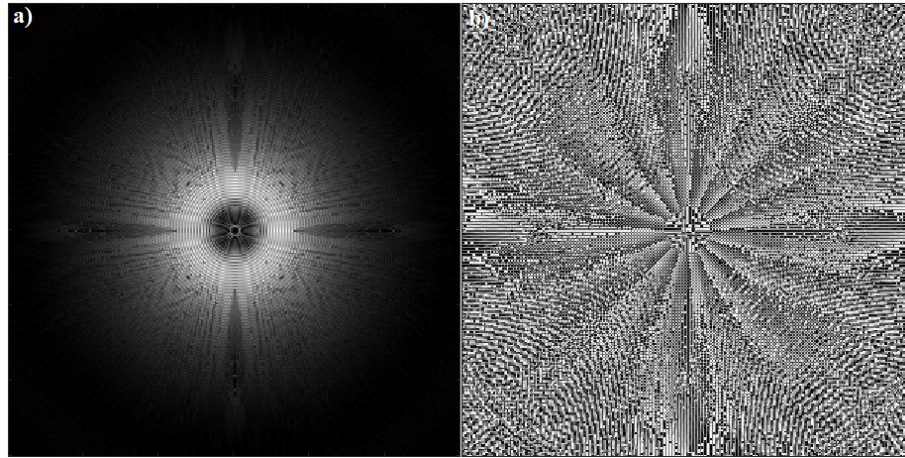


Figura 4.27: Recuperación de a) amplitud y b) fase de un VPC para 5 iteraciones.

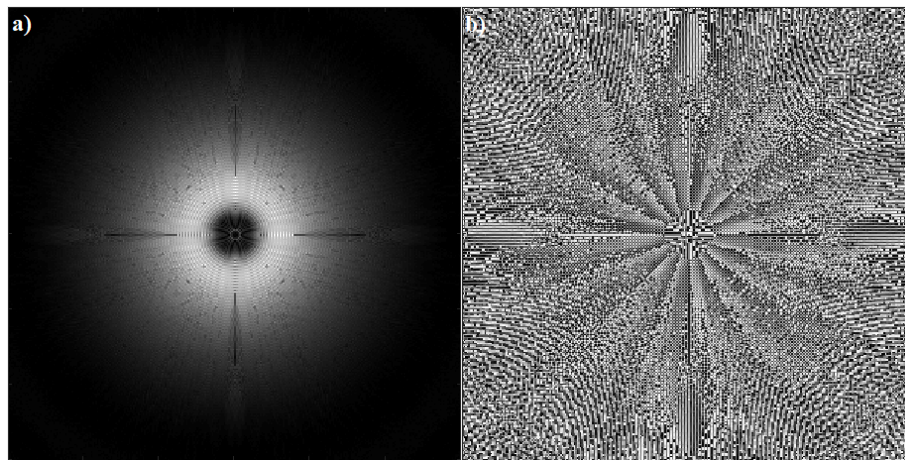


Figura 4.28: Recuperación de a) amplitud y b) fase de un VPC para 8 iteraciones.

4.3. Error cuadrático medio

El error puede ser monitoreado para cada iteración implementando la Ec. 3.37 dentro del algoritmo de Gerchberg-Saxton. En el caso de vórtice óptico perfecto, al utilizar una fase cuadrática como primera aproximación, dicho error disminuye de manera rápida dentro de las primeras 70 iteraciones hasta llegar a un valor al rededor de 1.5, tal como se ilustra en la (Fig. 4.39-a). El error para las iteraciones 70 a 140 disminuye de manera gradual (Fig. 4.39-b) y para la iteración número 200, el error calculado es alrededor de 0.0005 (Fig 4.39-c).

Al utilizar como primera aproximación a la fase del vórtice óptico perfecto una fase helicoidal, se observa que la disminución del error ocurre de manera más lenta que en el caso anterior, obteniendo un valor de 0.1 para las primeras 150 iteraciones (Fig. 4.40-a) y llegando a un valor de 0.004 para la iteración 200 (Fig. 4.40-b).

Para el vórtice parcialmente coherente, como podemos observar en la Fig. 4.41-a, el error disminuye con una rapidez significativa en las primeras iteraciones llegando a un valor de menos de 1 para la iteración 50 (Fig. 4.41-b). Para la iteración número 150 el error es de 0.039 (Fig. 4.41-c).

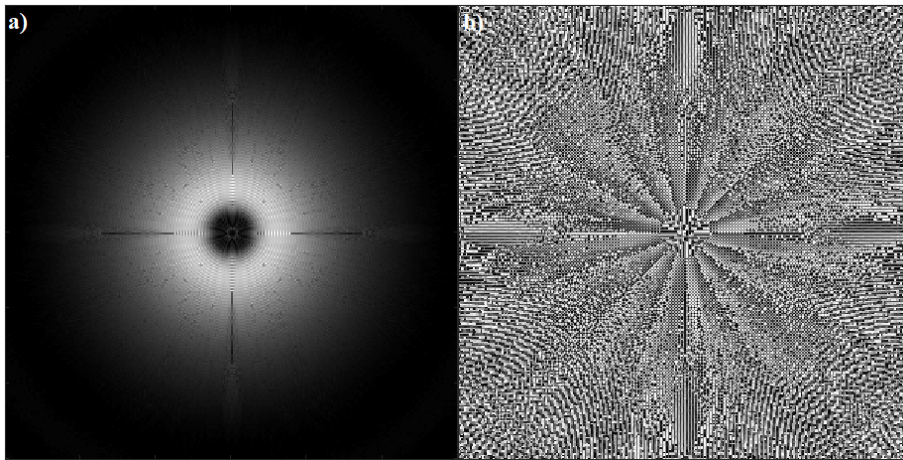


Figura 4.29: Recuperación de a) amplitud y b) fase de un VPC para 10 iteraciones.

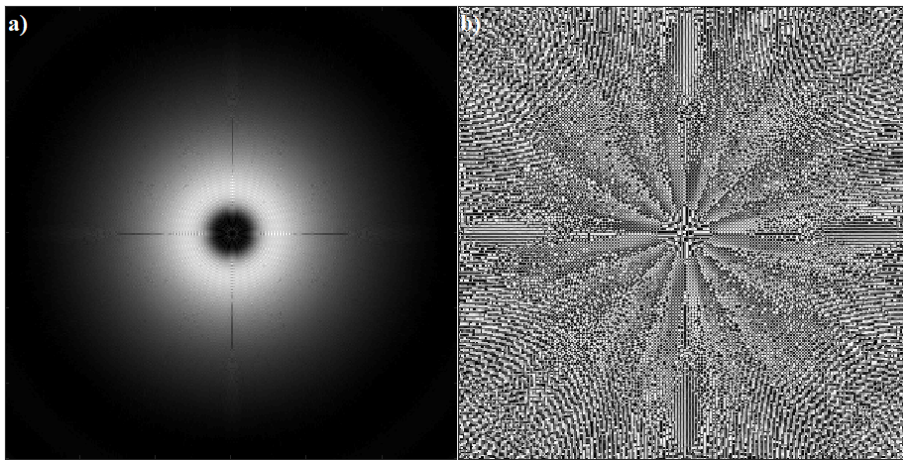


Figura 4.30: Recuperación de a) amplitud y b) fase de un VPC para 15 iteraciones.

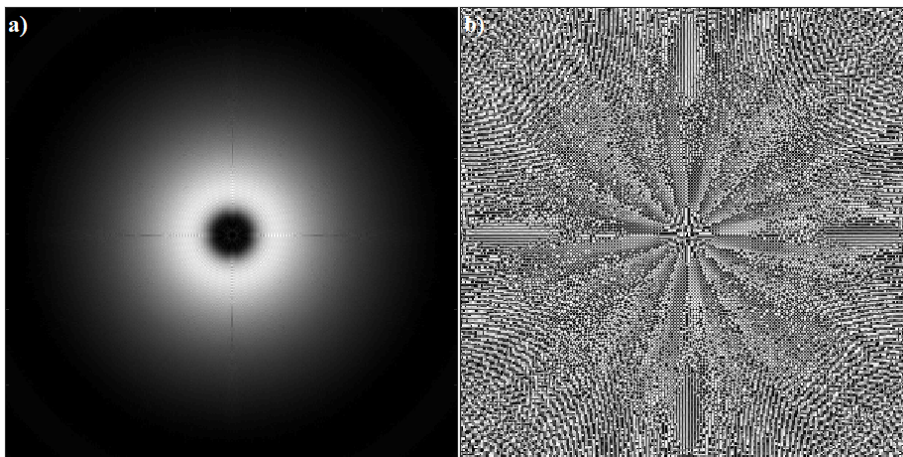


Figura 4.31: Recuperación de a) amplitud y b) fase de un VPC para 20 iteraciones.

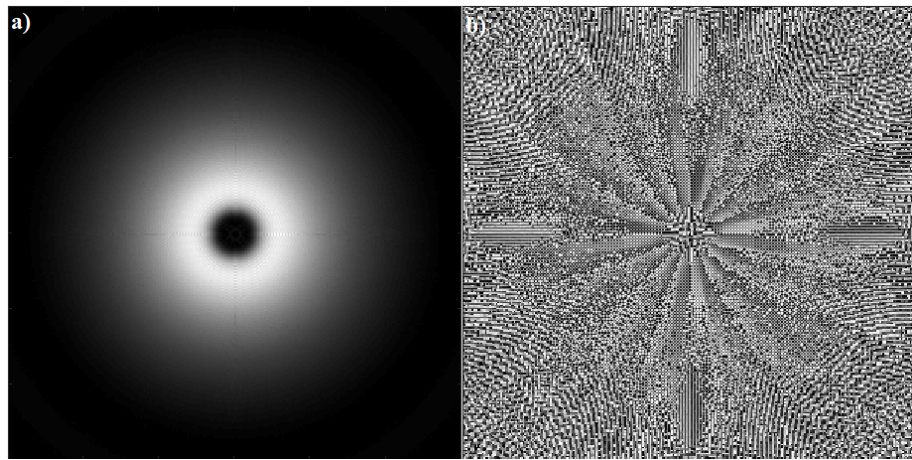


Figura 4.32: Recuperación de a) amplitud y b) fase de un VPC para 25 iteraciones.

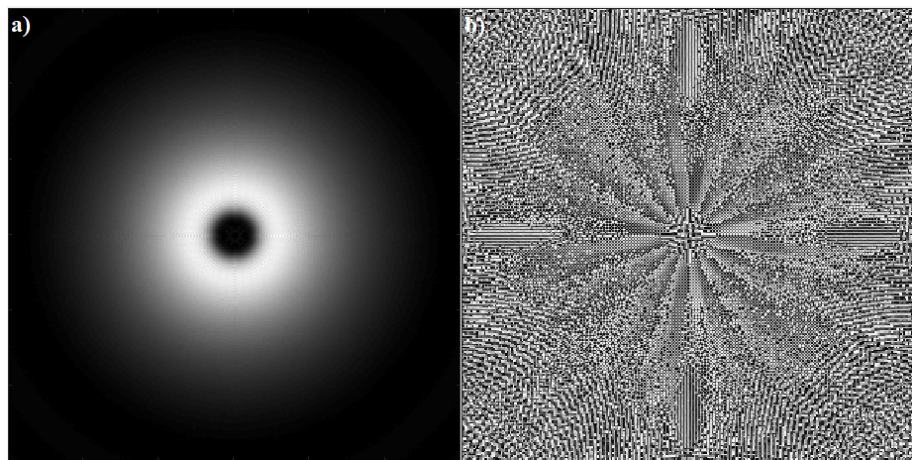


Figura 4.33: Recuperación de a) amplitud y b) fase de un VPC para 28 iteraciones.

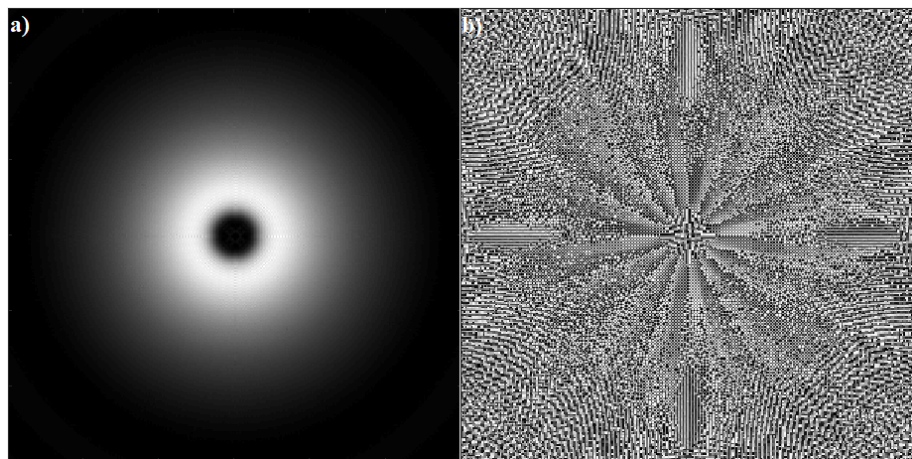


Figura 4.34: Recuperación de a) amplitud y b) fase de un VPC para 30 iteraciones.

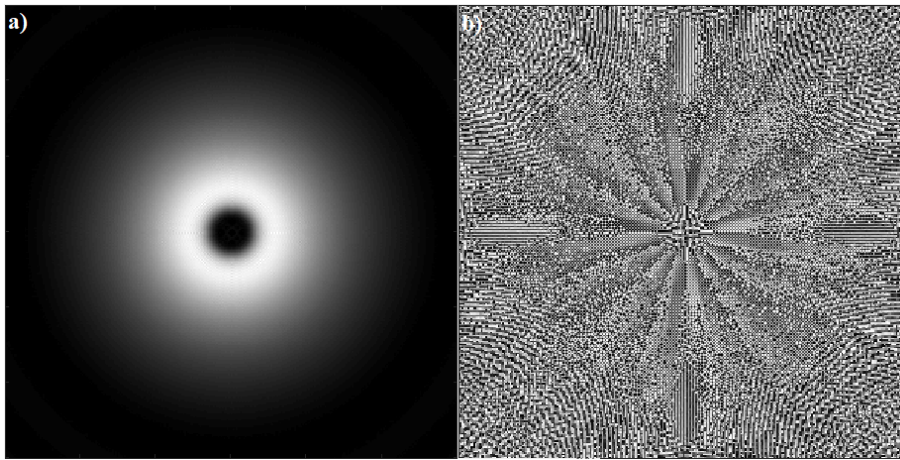


Figura 4.35: Recuperación de a) amplitud y b) fase de un VPC para 35 iteraciones.

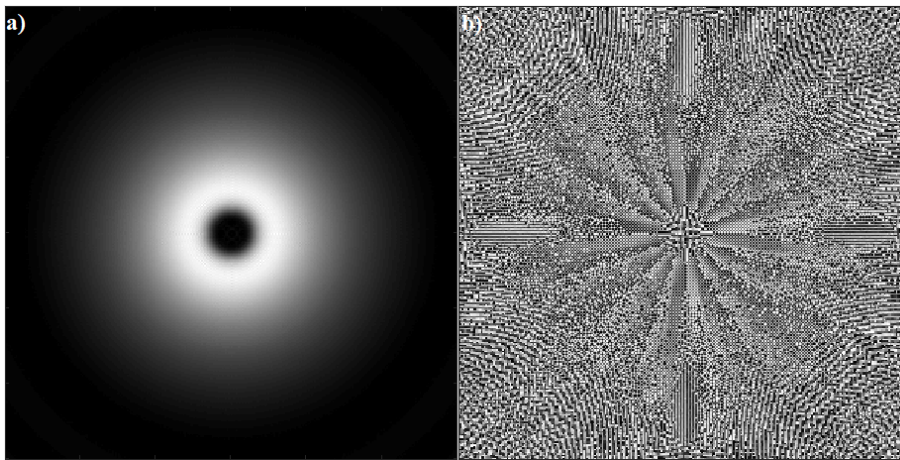


Figura 4.36: Recuperación de a) amplitud y b) fase de un VPC para 40 iteraciones.

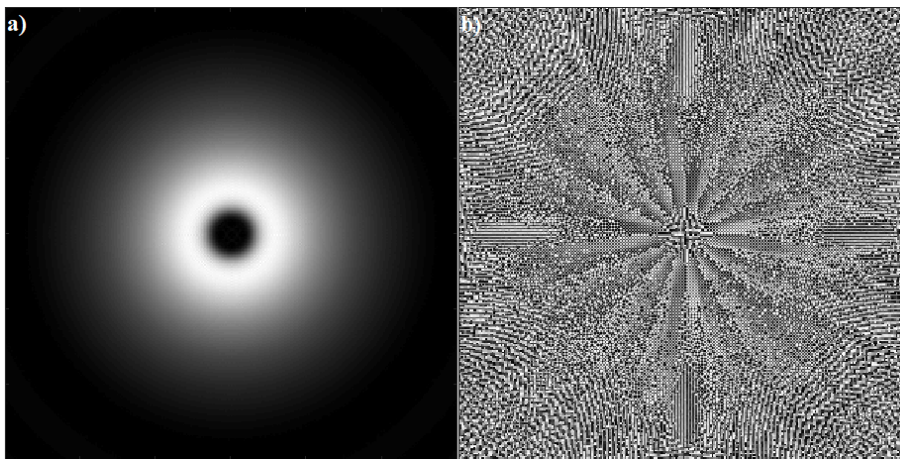


Figura 4.37: Recuperación de a) amplitud y b) fase de un VPC para 50 iteraciones.

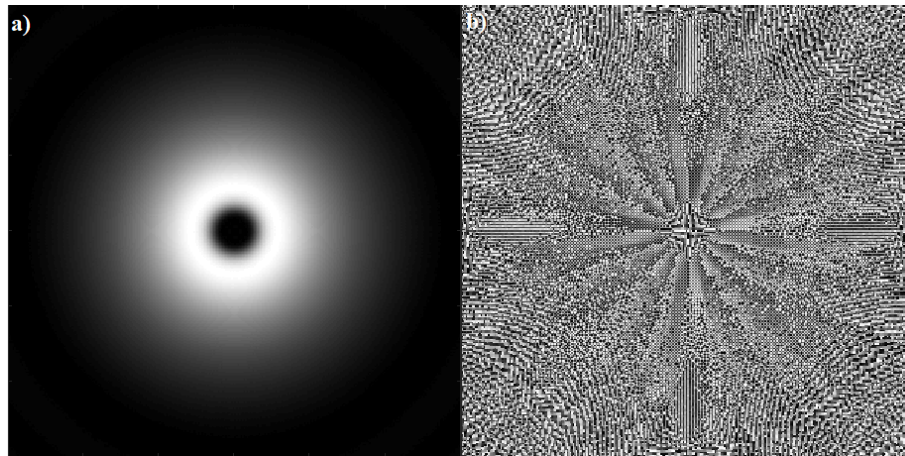


Figura 4.38: Recuperación de a) amplitud y b) fase de un VPC para 150 iteraciones.

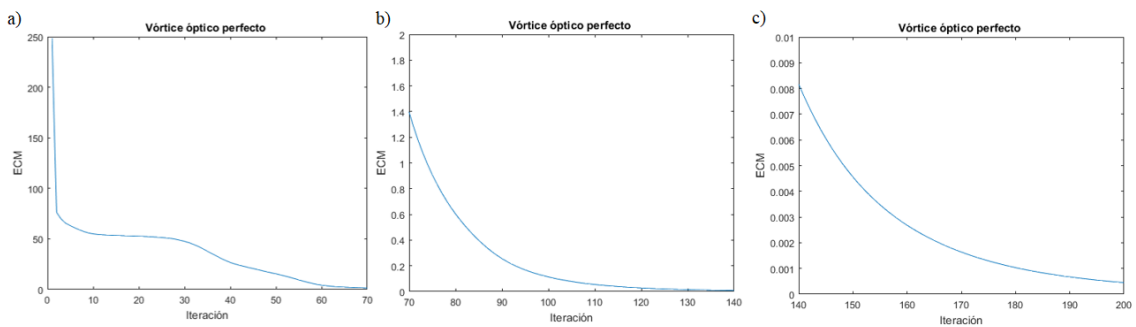


Figura 4.39: Error cuadrático medio obtenido entre la intensidad teórica y la estimada del VOP para a) las primeras 70 iteraciones, b) de la iteración número 70 a la 140 y c) las últimas 60 iteraciones utilizando una fase cuadrática como primera aproximación.

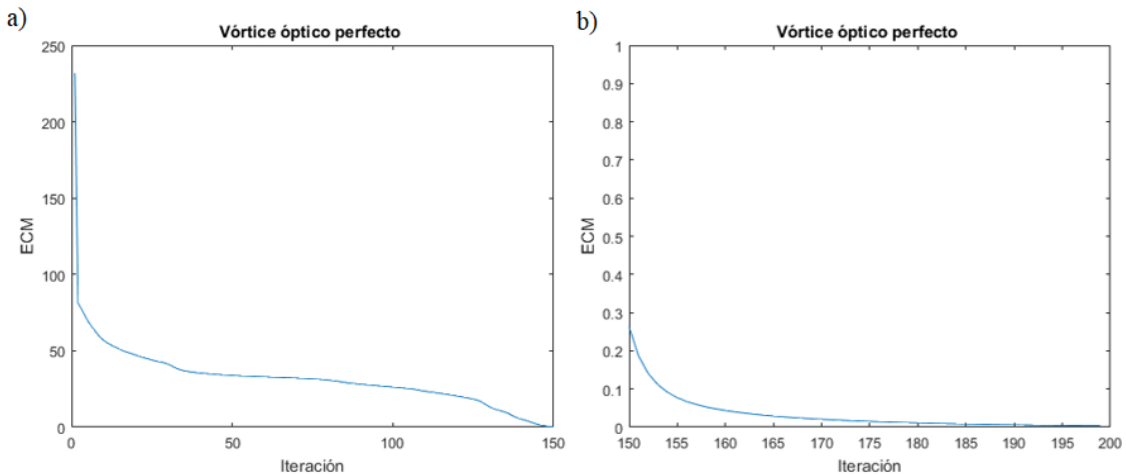


Figura 4.40: Error cuadrático medio obtenido entre la intensidad teórica y la estimada del VOP para a) las primeras 150 iteraciones y b) las últimas 50 iteraciones utilizando una fase helicoidal como primera aproximación.

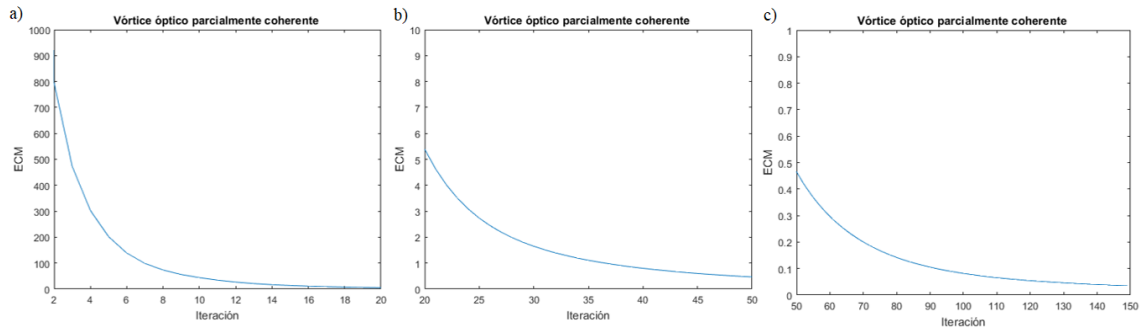


Figura 4.41: Error cuadrático medio obtenido entre la intensidad teórica y la estimada del VPC para a) las primeras 20 iteraciones, b) de las iteraciones número 20 a 50 y c) las últimas 100 iteraciones.

Capítulo 5

Conclusiones

Los resultados obtenidos indican que el algoritmo implementado es capaz de devolver una fase que genera los vórtices ópticos planteados. En el caso de un vórtice óptico perfecto, la reconstrucción ocurre con mayor rapidez al utilizar una fase cuadrática en lugar de una fase helicoidal, un hecho importante ya que indica que no es necesario tener indicios de la fase real del vórtice para obtener una buena aproximación a esta. No obstante, puede verse que ambas fases permiten reconstruir el vórtice perfecto con las restricciones impuestas en ambos planos. En el caso del vórtice óptico parcialmente coherente, la reconstrucción ocurre en un menor número de iteraciones que en el caso anterior. La importancia de lograr una buena aproximación a la fase que genera este tipo de vórtices aún sin tener algún conocimiento sobre la fase real radica en el hecho de que la realización física de un vórtice óptico parcialmente coherente requiere, tal como se mencionó, del diseño e implementación física de ensamblajes estadísticos lo que a su vez implica esfuerzos computacionales significativos. El algoritmo propuesto ahorra muchos de estos esfuerzos.

Se mencionó también que la fase obtenida por medio del algoritmo se encuentra envuelta, por lo que además es necesario utilizar algún método para su desenvolvimiento, sin embargo, para este trabajo no fue necesario obtener la fase desenvuelta ya que ésta se emplea directamente dentro del algoritmo y el objetivo radica en reconstruir la intensidad con la fase estimada.

La generación de vórtices ópticos se ha realizado a la fecha por medio de métodos que involucran sistemas ópticos elaborados y en muchos casos su implementación es complicada. En este trabajo hemos mostrado que es posible generarlos de forma directa a partir de un sistema de transformada de Fourier empleando el algoritmo iterativo de Gerchberg-Saxton. Hemos mostrado que a pesar de que la convergencia del algoritmo no está garantizada, en nuestro caso, éste devuelve la fase correcta, es decir la que teóricamente se requiere. En el caso de los vórtices ópticos esto es muy importante ya que la fase que los genera está estrechamente relacionada a la portación de momento angular orbital que tiene importantes aplicaciones en sistemas de atrapamiento y micromanipulación óptica.

Hasta donde sabemos la generación de vórtices por el mecanismo iterativo estudiado no había sido reportada hasta ahora. Por último, consideramos que los resultados obtenidos pueden tener un impacto importante en las técnicas de generación de vórtices ópticos conocidas y sus aplicaciones. En este sentido el método propuesto evita el uso de sistemas experimentales complicados y permite estudiar de forma práctica e inmediata las propiedades e implicaciones de diferentes clases de vórtices ópticos.

Bibliografía

- [1] X. WANG, Z. NIE, Y. LIANG, J. WANG, T. LI, B. JIA. *Recent advances on optical vortex generation*, Nanophotonics, **9**, pp. 1533–1556, (2018).
- [2] A. M. YAO AND M. J. PADGETT. *Orbital angular momentum: origins, behaviour and applications*, Advances in optics and photonics, **3**, pp. 161-204, (2011).
- [3] KOTLYAR VV, KOVALEV AA, SKIDANOV RV, MOISEEV OY, SOIFER VA. *Diffraction of a finite-radius plane wave and a Gaussian beam by a helical axicon and a spiral phase plate*, Journal of the Optical Society of America A. **24(7)**, pp. 1955–1964, (2007).
- [4] CHEN M, MAZILU M, ARITA Y, WRIGHT EM, DHOLAKIA K. *Dynamics of microparticles trapped in a perfect vortex beam*, Optics Letters **38(22)**, pp. 4919–4922, (2013).
- [5] GARCÍA-GARCÍA J, RICKENSTORFF-PARRAO C, RAMOS-GARCÍA R, ARRIZÓN V, OSTROVSKY A. *Simple technique for generating the perfect optical vortex.*, Optics Letters. **39(18)**, pp. 5305–5308, (2014).
- [6] A. S. OSTROVSKY, J. GARCÍA-GARCÍA, C. RICKENSTORFF-PARRAO, M. A. OLVERA-SANTAMARÍA. *Partially coherent diffraction-free vortex beams with a Bessel-mode structure*. Optics Letters, **42**, pp. 5182-5185 (2017).
- [7] GARRY TAYLOR. *The phase problem*, Acta Cryst., **59**, pp. 1881-1890 (2003).
- [8] Y. SHECHTMAN, Y. C. ELДАР, O. COHEN, H. N. CHAPMAN, J. MIAO, AND M. SEGEV. *Phase retrieval with application to optical imaging*, IEEE Signal Processing Mag., **32**, pp. 87-109, (2015).
- [9] HEINZ H. BAUSCHKE, PATRICK I. COMBETTES, D. RUSSELL LUKE. *Hybrid projection-reflection method for phase retrieval*, J. Opt. Soc. Am. A., **20**, pp. 1020 (2003).
- [10] HEINZ H. BAUSCHKE, PATRICK I. COMBETTES, D. RUSSELL LUKE. *Phase retrieval, error reduction algorithm, and Fienup variants: a view from convex optimization*, J. Opt. Soc. Am. A., **19**, pp. 1334 (2002).
- [11] J. GOODMAN. *Introduction to Fourier Optics*, Second Edition, McGraw Hill, (1996).
- [12] J. M. HERRERA FERNÁNDEZ. *Técnicas de síntesis y aplicaciones de elementos ópticos difractivos*. (2015)
- [13] E. HETCH. *Óptica*, Tercera Edición, Addison Wesley, (2000).
- [14] BORN, M. AND WOLF, E.. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light.*, Séptima edición, Cambridge University Press, (1999).
- [15] SOMMERFELD, A.. *Mathematische theorie der diffraction.*, Mathematische Annalen, **47(2)**, pp. 317–374 (1896).

- [16] P. L. COMBETTES. *The foundations of set theoretic estimation*. Proc. IEEE, **81**, pp. 182-208 (1993).
- [17] C. D. ALIPRANTIS, K. C. BORDER. *Infinite-Dimensional Analysis*. Segunda edición, Springer-Verlag, Berlin. (1999).
- [18] D. C. YOULA AND H. WEBB. *Image Restoration by the Method of Convex Projections: Part 1 Theory*. IEEE Transactions on Medical Imaging, **1**, pp. 81-94, (1982).
- [19] R. T. ROCKAFELLAR AND R. J.-B. WETS. *Variational Analysis*. Segunda edición, Springer-Verlag, Berlin. (1998).
- [20] C. TISSERON. *Notions de Topologie—Introduction aux Espaces Fonctionnels*. Segunda edición, Hermann, Paris, (1985).
- [21] D. C. YOULA. *Generalized image restoration by the method of alternating orthogonal projections*. IEEE Trans. Circuits Syst, **CAS-25**, pp. 694–702 (1978).
- [22] A. PAPOULIS. *A new algorithm in spectral analysis and band-limited extrapolation*. IEEE Trans. Circuits Syst, **CAS-22**, pp. 735–742 (1975).
- [23] D. VOELZ. *Computational Fourier Optics a MATLAB Tutorial*, SPIE Press, Washington. (2011).
- [24] JHON H. KARL. *An Introduction to Digital Signal Processing*, Academic Press, (1989).
- [25] R. W. GERCHBERG, W. O. SAXTON. *A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures*, Optik, **35**, pp. 237-246, (1972).
- [26] J. R. FIENUP. *Phase retrieval algorithms: a comparison*. Applied Optics, **21**, pp. pp. 2758-2769, (1982).
- [27] L. ALLEN, M. V. PADGETT AND M. BABIKER. *The orbital angular momentum of light*. Progress in Optics, **39**, pp. 291-372, (1999).
- [28] YIJE SHEN, XUEJIAO WANG, ZHENWEI XIE, CHANGJUN MIN, XING FU, QIANG LIU, MALI GONG, XIAOCONG YUAN. *Optical vortices 30 years on: OAM manipulation from topological charge to multiple singularities* Light: Science and Applications, (2019).
- [29] ANDREY S. OSTROVSKY, CAROLINA RICKENSTORFF-PARRAO, VÍCTOR ARRIZÓN. *Generation of the “perfect” optical vortex using a liquid-crystal spatial light modulator*. Optics Letters, **38**, pp. 534-536 (2013).
- [30] G. B. ARFKEN AND H. J. WEBER. *Mathematical Methods for Physicists*. Harcourt/Academic (2001).
- [31] JOAQUÍN GARCÍA-GARCÍA, CAROLINA RICKENSTORFF-PARRAO, RUBÉN RAMOS-GARCÍA, VÍCTOR ARRIZÓN, ANDREY S. OSTROVSKY. *Simple technique for generating the perfect optical vortex*. Optics Letters, **39**, pp. 5305-5308 (2014).
- [32] A. S. OSTROVSKY, J. GARCÍA-GARCÍA, C. RICKENSTORFF-PARRAO, M. A. OLVERA-SANTAMARÍA. *Partially coherent diffraction-free vortex beams with a Bessel-mode structure*. Optics Letters, **42**, pp. 5182-5185 (2017).
- [33] M.A. OLVERA SANTAMARÍA, J.L. RODRÍGUEZ GARCÍAPIÑA, J. GARCÍA GARCÍA, A.S. OSTROVSKY. *Generation of a partially coherent secondary source with Bessel-mode vortex structure by liquid crystal spatial light modulator*. Optics Communications, **439**, pp. 312-316 (2019).

- [34] L. MANDEL, E. WOLF. *Optical Coherence and Quantum Optics*. Cambridge University, (1995).