



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

CONSTRUCCIÓN DE UN SISTEMA TIPO SHG-FROG PARA LA
CARACTERIZACIÓN DE PULSOS LÁSER DE FEMTOSEGUNDOS

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

FÍSICO

P R E S E N T A :

FRANCO LUIS PRIEGO BOCHICCHIO

TUTOR

DR. JESÚS GARDUÑO MEJÍA



CIUDAD UNIVERSITARIA, CD. MX., 2022



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

*Para todos los que hicieron esto posible con su apoyo.
Nada que valga la pena hacer se hace sólo.*

Agradecimientos

Primero que nada agradezco a mi asesor, el Dr. Jesús Garduño, por su paciencia y por la ayuda que me dio a lo largo de este proceso. A los sinodales que forman parte del jurado por sus valiosas aportaciones que hicieron de este escrito su mejor versión. A los proyectos de la Dirección General de Asuntos del Personal Académico de la Universidad Nacional Autónoma de México PAPIIT-DGAPA, IN107821, así como al Consejo Nacional de Ciencia y Tecnología (CF-MG-2019-214961) y a Esperanza, Formación y Vida A.C.. De igual manera al Instituto de Ciencias Aplicadas y Tecnología por la beca que se me otorgó para la realización de este proyecto.

A mis compañeros en el ICAT y en el laboratorio de pulsos; en particular (pero no exclusivamente) Gus, Pablo, Itzel, Ramiro, Catalina y el M. en C. Carlos Román; que me enseñaron todas las herramientas que hicieron posible la parte experimental de este trabajo: el alineamiento del Mira, los detalles del uso de un cristal de SHG, impresión 3D, electrónica, entre mil otros trucos y habilidades que se usan diariamente en un laboratorio de óptica.

A todos los profesores que me inspiraron y me dirigieron a lo largo de mi trayectoria en la Facultad de Ciencias, que son más de los que podría nombrar aquí. A mis amigos fuera de la facultad, en particular Jesús, Adib y Luis, que me ayudaron a mantener un pie en la realidad. A mis amigos y compañeros en la Facultad sin quienes, no tengo duda, no sería posible que llegara hasta aquí; entre muchos otros puedo mencionar a Jose Luis, Ana, Sergio, Luis, pero en particular al resto de “los 4 fantásticos”: Silvestre, que completó los 4 y siempre hace que las discusiones, científicas y/o filosóficas, en el grupo sean interesantes.

Parra, que estuvo conmigo literalmente desde el primer día que puse pie en la facultad, si no hubiera tenido su amistad y su esfuerzo desde entonces nada de esto sería posible. A Lucero, mi compañera favorita de laboratorio, mi mejor amiga y luego mi primer y único amor. Sin tu compasión y tu amor no sería la persona que soy hoy, gracias por ayudarme a crecer a tu lado.

A mi mamá, Daniela, y mi nonna Gianna por enseñarme tanto durante mi vida, todo lo bueno que soy viene de ellas. Además por su gran paciencia y apoyo en los últimos años. A Bruno por haberme aguantado toda la vida y aún así ser infinitamente afable y bondadoso. A mis tíos, Tommaso y María Rosa, por su gran calidez y generosidad, y por ser grandes ejemplos a seguir. A toda mi familia por el impacto positivo que todos han tenido sobre mí, nunca olvido la suerte que tengo de tener una familia como esta.

Resumen

El objetivo de este trabajo fue el rediseño y modificación de un aparato que usa la técnica SHG-FROG (Second Harmonic Generation - Frequency Resolved Optical Gating) para la caracterización de pulsos ultracortos de láser. La técnica SHG-FROG utiliza un medio óptico no lineal y un algoritmo iterativo para caracterizar la intensidad y la fase de un pulso. Las mejoras principales al aparato respecto a su versión anterior tienen el objetivo de aumentar la velocidad de adquisición, es decir, que pueda hacer más mediciones en un menor tiempo. Otras consideraciones que se tomaron en cuenta para el rediseño fueron la versatilidad del aparato, que sea fácil de modificar y mejorar, y que sea fácil de usar. Específicamente, se rediseñó y construyó la parte encargada de la generación y adquisición del espectro que, en el diseño anterior, limitaba la velocidad de adquisición. De esta manera, el rediseño permitirá la reconstrucción de la fase y amplitud de pulsos ultracortos a frecuencias de unos cuantos Hertz. También se escribió un software que analiza la traza resultante de la medición para obtener la caracterización cuantitativa de los pulsos medidos mediante el algoritmo FROG por PCGP (Principal Component Generalized Projections).

Índice general

Agradecimientos	II
Resumen	IV
Introducción general	XI
1. Pulsos ultracortos	1
1.1. Características de un pulso ultracorto	2
1.2. Fase temporal y fase espectral	5
2. Óptica no lineal	9
2.1. Polarización en el régimen lineal	10
2.2. Polarización en el régimen no lineal	11
2.3. SHG y suma de frecuencias	12
2.4. Phase matching	16

<i>ÍNDICE GENERAL</i>	VI
2.4.1. Phase matching Para SHG	17
3. La técnica FROG	21
3.1. La traza FROG	26
4. El algoritmo FROG	29
4.1. El algoritmo FROG básico	29
4.2. El algoritmo FROG por PCGP	33
5. Software en python	40
6. Diseño del dispositivo	49
6.1. El autocorrelador	50
6.1.1. La línea de retardo	55
6.2. El espectrómetro	60
6.2.1. Espejo rotatorio y cámara	62
6.3. Sistema de control	64
7. Pruebas al dispositivo	69
7.1. Servo SG90	71
7.2. Motor a pasos Nema 17 con micropasos	73

<i>ÍNDICE GENERAL</i>	VII
8. Conclusiones	77
A. Algunos cálculos	80
B. Información práctica sobre el software de análisis	83

Índice de figuras

1.1. Representaciones de un pulso ultracorto.	4
2.1. Señal distorsionada por “clipping”	12
2.2. Diagrama de la generación de segundo armónico (SHG)	14
2.3. Diagrama de la suma de frecuencias	15
3.1. Diagrama de la autocorrelación de una señal.	22
3.2. Diagrama de un autocorrelador de intensidad por SHG.	23
3.3. Diagrama de un arreglo SHG-FROG	25
3.4. Dos ejemplos de trazas SHG-FROG	26
4.1. Diagrama de la estructura general de un algoritmo FROG genérico	31
4.2. Diagrama del concepto detrás del funcionamiento de un algoritmo de proyecciones generalizadas	34
4.3. Diagrama de la estructura de una iteración del algoritmo FROG por PCGP	38

5.1. Trazas para un pulso con perfil de intensidad gaussiano y fase constante . . .	43
5.2. Amplitud y fase del pulso simulado y recuperado, un pulso gaussiano con fase constante	43
5.3. Trazas para un pulso con perfil de intensidad gaussiano y fase cuadrática .	44
5.4. Amplitud y fase del pulso simulado y recuperado, un pulso gaussiano con fase cuadrática	44
5.5. Trazas para un pulso con perfil de intensidad bimodal con fase cuadrática .	45
5.6. Amplitud y fase del pulso simulado y recuperado, un pulso con dos picos gaussianos, cada uno con fase cuadrática	45
5.7. Trazas para un tren de dos pulsos gaussianos con fases cuadráticas de distinto signo	46
5.8. Amplitud y fase del pulso simulado y recuperado, tren de dos pulsos gaussianos con fases cuadráticas de signos opuestos	46
5.9. Diagrama de flujo que representa el funcionamiento del script de análisis. .	48
6.1. Esquema de la formación de una traza	50
6.2. Foto del arreglo durante la alineación	51
6.3. Foto de la sección inicial del autocorrelador	53
6.4. Modelo 3D de la pieza que sostiene la montura del cristal no lineal.	54
6.5. Prueba de la línea de retardo mediante un interferómetro	56
6.6. Prueba de la línea de retardo mediante un interferómetro, escala más cercana	56

6.7. Prueba de la línea de retardo mediante un interferómetro, oscilación de la línea	57
6.8. Gráfica: Movimiento de la línea de retardo contra el ciclo de trabajo	59
6.9. Modelo 3d de la pieza diseñada para sostener el espejo rotatorio y la cámara	61
6.10. Pieza impresa en 3D montada en el arreglo	61
6.11. Comparación entre dos espectros tomados con la cámara	63
6.12. Circuito en PCB que forma parte del sistema de control.	66
6.13. Diagrama de la parte del circuito que controla la línea de retardo	66
7.1. Resultado de un barrido de un espectro constante con el servomotor.	71
7.2. Resultado de un barrido de un espectro constante con el motor a pasos.	73
7.3. Gráfica del marginal de retardo de la imagen 7.2.	73
7.4. izquierda: Gráfica de la posición relativa del espectro para cada paso. Derecha: Imagen compilada con cada posición del espectro	75

Introducción general

En esta tesis se presentará el rediseño y construcción de un dispositivo para la caracterización de la amplitud y fase de pulsos láser ultracortos (con duraciones en el orden de femtosegundos ^[1]) usando la técnica FROG (Frequency Resolved Optical Gating) por SHG (Second Harmonic Generation). Se describirá la técnica, así como sus fundamentos teóricos, las consideraciones prácticas que se deben tomar para llevarla a cabo, y los métodos de análisis necesarios para cuantificar sus resultados.

Desde la creación de tecnologías como el *Q-switching* y el *Mode-Locking*, o *Amarre de Modos*, que han permitido la generación de pulsos láser de duración cada vez más corta, los láseres pulsados han sido usados para una gran variedad de aplicaciones. Las principales características de los pulsos láser generados con estas técnicas son su corta duración y las altas potencias pico que se pueden lograr. En el caso de pulsos con duración de femtosegundos generados por la técnica de amarre de modos por efecto Kerr óptico, se pueden generar potencias pico de cientos de kW [Lemoff and Barty, 1992], y en sistemas amplificados por medio de la técnica CPA (Chirped Pulse Amplification) se logran potencias hasta del orden de TW [Maine et al., 1988].

Entre muchas otras aplicaciones, uno de sus principales usos es la espectroscopía ultrarrápida, donde se usan pulsos para hacer mediciones espectroscópicas de estructuras o procesos con corta duración, de manera análoga a como un flash de alta velocidad se usa para tomar fotografías de eventos cortos. Las altas potencias pico que pueden tener los pulsos cortos hacen que sean útiles para aplicaciones de tipo industrial (cortar o perfo-

^[1]1fs = 10⁻¹⁵s

rar ciertos materiales), o incluso para generar y estudiar plasmas. Tienen usos en campos como la astrofísica, donde se usan para generar “peines ópticos” para medir frecuencias con precisión, y la ingeniería de comunicaciones, donde se pueden usar para transmitir datos a velocidades de terabits por segundo, sin mencionar que son fundamentales para el estudio de la óptica no lineal [Weiner, 2009]. Para todos estos usos es fundamental conocer la estructura, o tan siquiera la duración, de los pulsos que se usarán.

Para medir la duración de cualquier evento, es necesario tener un evento más corto con el cual comparar, así como para medir una longitud se usa una regla con marcas de menor tamaño a lo que se pretende medir. Esto representa un problema cuando se quieren medir pulsos ultracortos, que llegan a tener duraciones de femtosegundos. De hecho son los eventos mas cortos que se han creado en un laboratorio, ¿Cómo se pueden medir sin un evento más corto con el cual comparar? Es claro entonces que en el caso de pulsos de femtosegundos, no es posible medir directamente la duración con detectores electrónicos, dado que tienen un tiempo de respuesta muy largo en comparación a la duración del pulso, y mucho menos medir su fase, que es uno de los principales retos en la caracterización óptica.

La técnica FROG fue la primera técnica capaz de medir la amplitud y la fase de los pulsos ultracortos. Es una técnica que combina un diseño óptico relativamente simple, con un poderoso algoritmo iterativo para recuperar la amplitud y fase de un pulso. El diseño óptico se basa en un arreglo interferométrico y espectrométrico, y hace uso de la generación de efectos ópticos no lineales. Desde su origen en 1991 [Trebino, 2000] ha sido un estándar en el campo de los pulsos, el FROG y sus variaciones todavía constituyen la vasta mayoría de las técnicas que se usan en la medición de pulsos ultracortos. Dado que la mayoría de las alternativas no son más que variaciones de FROG ^[2], la técnica original sigue teniendo vigencia en los laboratorios alrededor del mundo por su sencillez y versatilidad, que además la hacen particularmente útil cuando se busca rapidez en la medición.

^[2]Existen excepciones como GRENOUILLE [Trebino, 2000], que es una técnica en principio “más sencilla” porque no tiene partes móviles, pero que requiere elementos ópticos caros y poco comunes

En el primer capítulo se presenta una introducción general a los pulsos láser. Se describen sus características, como lo son sus perfiles de intensidad y fase, que se buscan medir mediante la técnica FROG. Se incluye también una discusión sobre la duración de un pulso y las distintas maneras de cuantificarlo.

El segundo capítulo comienza con una introducción a la óptica no lineal. Se describen con especial atención los efectos no lineales de segundo orden. Es de particular importancia práctica la discusión sobre el *phase matching*. Estos dos primeros capítulos forman la base teórica para el resto del trabajo.

En el tercer capítulo se describen los fundamentos de la técnica FROG. Se presenta la autocorrelación no lineal como base para introducir la técnica FROG, y se describe el arreglo óptico típico para un sistema SHG-FROG como el que se construyó para este trabajo, así como las expresiones matemáticas relevantes. Además se presenta la traza FROG y sus características.

En el cuarto capítulo se describe el algoritmo iterativo que se utiliza para recuperar las características cuantitativas del pulso a partir de la traza FROG. Se describen primero las generalidades de los algoritmos de análisis comúnmente usados, y posteriormente se discuten los detalles de la versión específica del algoritmo que se usó en este trabajo, conocida como *algoritmo de proyecciones generalizadas por componentes principales* o PCGP (por sus siglas en inglés: Principal Component Generalized Projection).

En el quinto capítulo se presenta la implementación del algoritmo FROG-PCGP que se creó para este trabajo. Se discuten algunas consideraciones prácticas de su uso y se muestran los resultados de varias pruebas de su funcionamiento realizadas con trazas simuladas. Un análisis y descripción más detallada se encuentra en el anexo B.

En el sexto capítulo se muestra y se describe el diseño del dispositivo FROG construido. Se describen cada una de sus secciones y se discuten los cambios que se efectuaron en cada una y las decisiones que llevaron al diseño final. También se describen algunas mediciones que se realizaron durante el proceso de diseño.

En el séptimo capítulo se presentan pruebas del funcionamiento del dispositivo, que conciernen en particular a la medición y despliegue del espectro, realizadas con un espejo giratorio, lo que representa la modificación más drástica al diseño del dispositivo.

Capítulo 1

Pulsos ultracortos

Un pulso láser es una emisión de energía electromagnética repentina y de poca duración. Como cualquier onda de luz, consiste en una oscilación de los campos eléctrico y magnético. En el caso de un pulso, a diferencia del caso más sencillo y más conocido de una onda continua, esa oscilación está limitada a un intervalo de tiempo específico^[1], y por tanto se suele representar como una función oscilatoria multiplicada por una función envolvente. No hay una definición específica de qué duración debe tener un pulso para considerarse “ultracorto”, pero se le suele llamar ultracortos a los pulsos cuando tienen una duración de menos de unas decenas de picosegundos [Trebino, 2000] [RP-Photonics, sf]. En una sección más adelante se discutirá más sobre cómo se define la duración de un pulso.

Cuando se considera un pulso de luz muy corto surgen fenómenos interesantes. Uno de los más importantes es que la duración de un pulso corto resulta estar relacionada con su ancho espectral [Trebino, 2000]. Se sabe en el mundo de la óptica y de las señales en general, que solo una onda continua e infinita tiene una frecuencia única, cualquier señal limitada en el tiempo debe tener un ancho espectral, entre más esté limitada mayor será ese ancho espectral ^[2]. Cuando se trata de una señal tan limitada en el tiempo como lo es

^[1]Cabe mencionar que en la realidad toda emisión electromagnética está limitada a un intervalo de tiempo, como todo en este mundo tienen inicio y fin, sin embargo las ondas continuas se suelen describir matemáticamente como ondas continuas idealizadas, infinitas en el tiempo.

^[2]En el caso de la óptica, este hecho se puede explicar desde la perspectiva de la cuántica con el principio

un pulso ultracorto, ese ancho espectral relacionado a la duración se vuelve el factor más importante en el espectro del pulso. Esto podría sugerir una manera de decir algo sobre la duración de un pulso, y de hecho el ancho espectral de un pulso puede dar una cota inferior a su duración, pero no una medida de su duración real.

1.1. Características de un pulso ultracorto

Dado que en este trabajo son de interés las características temporales de los pulsos y no las espaciales, se describirán como su campo eléctrico $\mathcal{E}(t)$ en función sólo del tiempo sin considerar su dependencia de variables espaciales. Más aún, se usará lo que se conoce como la *aproximación escalar*, es decir, se considerará que el campo tiene polarización lineal y se escribirá sólo su magnitud para ahorrar el tener vectores por todos lados y tratar con todas sus componentes.

El campo eléctrico de un pulso se suele representar como un producto de una función que representa la amplitud (función del tiempo), y una función que representa la parte oscilatoria como se muestra en la ecuación 1.1.

$$\mathcal{E}(t) = \frac{1}{2} \sqrt{I(t)} e^{i\Gamma(t)} + c.c. \quad (1.1)$$

Donde $I(t)$ es la intensidad luminosa que se mediría, $\Gamma(t)$ es una función que describe la oscilación, y se añade $+c.c.$, es decir se suma el complejo conjugado de lo anterior, para que $\mathcal{E}(t)$ sea real. Solo se describe el campo eléctrico, pues el campo magnético es igual salvo multiplicación por algunas constantes. Para simplificar las matemáticas de aquí en adelante se usará también lo que se conoce como la aproximación de señal analítica, se omitirá el complejo conjugado. De esta manera la función será compleja, así que no representa el valor del campo como tal, pero contendrá toda la misma información.

En casi todos los casos de interés, el pulso se comporta de manera que su espectro de incertidumbre; los fotones que componen un pulso muy corto están restringidos en el tiempo a la duración del pulso corto, y por lo tanto su energía (y equivalentemente su frecuencia) debe estar sumamente indefinida, y eso aumenta el ancho espectral del pulso

tiene una frecuencia central o promedio ω_0 y los valores del espectro caen en un intervalo $\Delta\omega$, pequeño comparado con ω_0 , alrededor de ella. En ese caso se puede escribir la parte oscilatoria como una oscilación con la frecuencia central, y describir las variaciones en la frecuencia usando un término de fase dependiente del tiempo [Diels and Rudolph, 2011], de la manera que se muestra en la ecuación 1.2. [Trebino, 2000]

$$\mathcal{E}(t) = \frac{1}{2} \sqrt{I(t)} e^{i(\omega_0 t - \phi(t))} + c.c. \quad (1.2)$$

A $\phi(t)$ se le conoce como la *fase temporal* del pulso. La parte oscilatoria que lleva la frecuencia central, $e^{i\omega_0 t}$, se conoce como la *onda portadora* [Trebino, 2000]. Si de la expresión en la ecuación 1.2 se quita la parte de la onda portadora, queda lo que se define como la *amplitud compleja*, dada en la ecuación 1.3.

$$E(t) = \sqrt{I(t)} e^{-i\phi(t)} \quad (1.3)$$

La amplitud compleja es esencialmente la cantidad que se busca encontrar cuando se habla de caracterizar un pulso, ya que contiene el perfil de intensidad y la fase del pulso. En la figura 1.1 se muestra gráficamente la relación entre el campo eléctrico como tal, la magnitud de la amplitud compleja, y la intensidad del pulso. Considerando esta definición y las aproximaciones que ya se mencionaron, el pulso se puede escribir como se muestra en la ecuación 1.4

$$\mathcal{E}(t) = E(t) e^{i\omega_0 t} \quad (1.4)$$

El pulso también se puede representar en el espacio de frecuencias, simplemente haciendo una transformada de Fourier a $\mathcal{E}(t)$ de tal manera que:

$$\tilde{\mathcal{E}}(\omega) = \int_{-\infty}^{\infty} \mathcal{E}(t) e^{-i\omega t} dt, \quad (1.5)$$

y si se separa $\tilde{\mathcal{E}}(\omega)$ en un perfil de intensidad y una fase se puede escribir:

$$\tilde{\mathcal{E}}(\omega) = \sqrt{S(\omega)}e^{-i\varphi(\omega)}, \quad (1.6)$$

donde $S(\omega)$ es el espectro del pulso y $\varphi(\omega)$ es la *fase espectral* del pulso. En la siguiente sección se discutirá la relación entre las dos fases, temporal y espectral, y su significado.

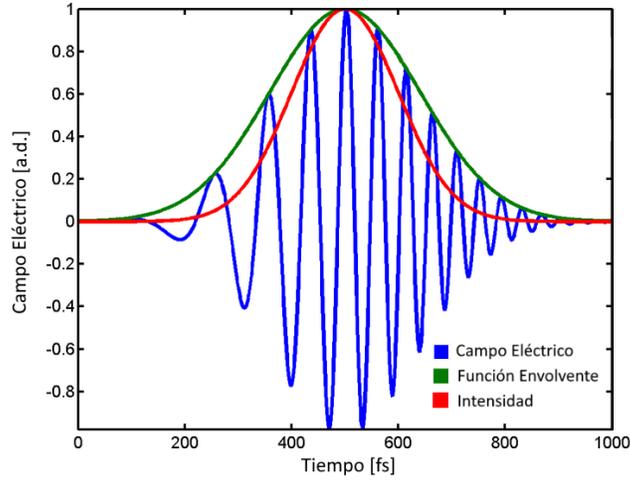


Figura 1.1: Gráfica adimensional de las representaciones de un pulso ultracorto. La curva azul representa el campo $\mathcal{E}(t)$, la curva verde representa la envolvente, la magnitud de la amplitud compleja dada en la ecuación 1.3, y la curva roja representa la intensidad del pulso.

Cuando se habla de pulsos cortos se suele hablar de la duración de un pulso. Existen varias maneras de describir la duración una vez que se conoce $I(t)$. La manera más sencilla es usar el FWHM ("full width at half maximum", o ancho completo a la mitad del máximo), que es el ancho de la curva medido a la mitad de su altura máxima. Otra medida similar que se usa comúnmente es el "Half Width $1/e$ " (notada $\tau_{HW1/e}$), que es la mitad del ancho de la curva a una altura igual a su altura máxima entre e [Trebino, 2000]. Ésta última medida es particularmente útil si el perfil de intensidad del pulso es gaussiano, pues el perfil se puede escribir fácilmente en términos de ella:

$$I(t) = I_0 e^{-(t/\tau_{HW1/e})^2}. \quad (1.7)$$

Los anchos FWHM y HW1/e son fáciles de obtener, sin embargo son menos útiles cuando se trata de pulsos con mucha estructura. Para esos casos una medida más apropiada es el ancho RMS (notado τ_{rms}), que es simplemente el segundo momento de la distribución $I(t)$ [Trebino, 2000]. El ancho RMS es útil para pulsos con estructura porque mientras que el ancho FWHM por ejemplo ignora cualquier estructura con menos de la mitad de la altura del pulso, el ancho RMS no ignora nada, y de hecho pesa más las estructuras lejanas al centro del pulso.

También en muchas situaciones es útil conocer lo que se conoce como el *producto tiempo-ancho de banda* (TBP). El TBP, como su nombre lo indica es el producto del ancho temporal y el ancho espectral de un pulso, ambos medidos con cualquiera de las definiciones anteriormente mencionadas. El TBP funciona como una medida de la complejidad de un pulso, entre más pequeño es, más “simple” el pulso en cuestión [Trebino, 2000].

1.2. Fase temporal y fase espectral

La fase temporal $\phi(t)$ es la manera en la que se denotarán cambios de frecuencia, se puede ver que la oscilación del pulso en la ecuación 1.2 será afectada por $\phi(t)$. De hecho la frecuencia del pulso en un tiempo particular, conocida como la frecuencia instantánea $\omega_{ins}(t)$ se puede escribir como [Trebino, 2000] [Diels and Rudolph, 2011]:

$$\omega_{ins} \equiv \omega_0 - \frac{d\phi}{dt}. \quad (1.8)$$

De manera análoga para la fase espectral, $\phi(\omega)$, se puede definir el retardo de grupo t_{grupo} (que representa el tiempo que se retarda cada componente espectral de la luz debido a la fase) como [Trebino, 2000]:

$$t_{grupo} \equiv \frac{d\phi}{d\omega}. \quad (1.9)$$

Suele ser útil escribir las fases como sus expansiones en serie de Taylor:

$$\phi(t) = \phi_0 + \phi_1 t + \phi_2 t^2 + \phi_3 t^3 + \phi_4 t^4 + \dots \quad (1.10)$$

$$\varphi(\omega) = \varphi_0 + \varphi_1 \omega + \varphi_2 \omega^2 + \varphi_3 \omega^3 + \varphi_4 \omega^4 + \dots \quad (1.11)$$

Donde ϕ_n (o φ_n) se conoce como la fase (temporal o espectral) de orden n . En un pulso simple o “bien portado”, solo serán relevantes los términos de órdenes bajos de cada una, sin embargo para un pulso más complejo pueden entrar en juego las fases de altos órdenes.

La fase de orden cero es igual en ambos espacios (temporal y espectral), es decir $\phi_0 = \varphi_0$. Representa la fase relativa entre las oscilaciones del campo eléctrico como tal y la envolvente. Viendo las ecuaciones 1.8 y 1.9 es claro que las fases de orden cero no influyen en la frecuencia instantánea ni en el retardo de grupo.

Las fases de orden uno corresponden a una traslación en el dominio opuesto, es decir, una fase temporal de orden uno corresponde a una traslación del espectro, y una fase espectral de orden uno corresponde a una traslación en el perfil de intensidad, en otras palabras un retraso. Ésto se puede ver fácilmente en las ecuaciones las ecuaciones 1.8 y 1.9, donde una fase lineal resulta en un término constante añadido a cada una.

La fase temporal de orden dos corresponde a lo que se conoce como *chirp lineal*. Cuando se tiene fase temporal de segundo orden la frecuencia instantánea se vuelve:

$$\omega_{ins} = \omega_0 - 2\phi_2 t, \quad (1.12)$$

y por lo tanto la frecuencia instantánea se vuelve una función lineal del tiempo, a diferentes tiempos la luz que compone el pulso contiene distintas frecuencias [Diels and Rudolph, 2011]. La frecuencia instantánea aumenta o disminuye conforme avanza el tiempo, dicho de otra manera, las frecuencias que componen el pulso pasan por una cierta posición en orden de mayor a menor, o de menor a mayor, dependiendo del signo de ϕ_2 . En la figura 1.1 el pulso tiene chirp positivo, es decir el segundo término de la ecuación 1.12 es positivo y por lo tanto las frecuencias bajas del pulso llegan antes (i.e. corresponden a tiempos

menores). La mayoría de los medios tienen dispersión tal que inducen un chirp positivo en un pulso cuando los atraviesa, retrasan las frecuencias mayores. Ésto incluye al aire, y por lo tanto se puede esperar que cualquier pulso generado eventualmente obtenga chirp, ya sea por los elementos ópticos que atraviesa o simplemente por propagarse en el aire. Al retrasarse ciertas frecuencias más que otras, se alarga la duración del pulso, y para pulsos suficientemente cortos puede ser un efecto de importancia mayor en su duración. Es por esto que para quien quiere generar o ver pulsos de corta duración es importante evitar o compensar el chirp.

Se puede mostrar que cuando se tiene una fase espectral de orden dos distinta de cero, también corresponde a chirp lineal [Trebino, 2000].

Cuando se tiene fase espectral de orden tres se genera lo que se conoce como chirp cuadrático. En éste caso el retardo de grupo es:

$$t_{grupo} = 3\varphi_3 t^2. \quad (1.13)$$

Ésto causa que la frecuencia central del pulso llegue antes (o después si φ_3 es negativo) que las frecuencias en las orillas de la distribución. Al llegar al mismo tiempo las frecuencias a cada lado de la central, interfieren y generan una oscilación en la intensidad. Ésto se ve como pequeños pulsos “satélite” enfrente (o detrás) de un pulso principal.

Así podría seguirse analizado el efecto de cada orden superior en el pulso, pero con esos primeros es suficiente para entender que las distorsiones de fase causadas por medios con dispersión de órdenes mayores puede tener efectos inesperados y de gran importancia en el pulso, y eso por esto que conocer la fase es tan importante como conocer el perfil de intensidad.

Cuando la fase espectral $\varphi(\omega)$ es constante^[3], se considera que el pulso está *limitado en ancho de banda*. En este caso la duración del pulso será la mínima que puede ser dado su ancho de banda, y por lo tanto, para una forma dada del pulso, el TBP será mínimo [Weiner,

^[3]o cuando es lineal, pues como ya se mencionó una variación lineal representa solo una traslación en el tiempo

2009]. Conforme aparezcan variaciones en la fase espectral, el TBP aumentará, y la duración del pulso no será la mínima posible para su ancho de banda.

Capítulo 2

Óptica no lineal

Después de la invención de los primeros láseres en la década de los 60's se empezaron a encontrar fenómenos ópticos que no se podían explicar desde el punto de vista de la óptica que existía hasta ese tiempo [Boyd, 2008] [Sutherland, 2003]. Ésto es lo que llevó a la creación del campo de la *óptica no lineal*, que se encarga de estudiar la interacción de la luz con la materia en el régimen de intensidad donde la respuesta de los materiales deja de ser proporcional a la amplitud de la luz en cuestión.

Cuando una onda electromagnética incide en un material dieléctrico, causa que las cargas dentro del material se muevan. Bajo la acción de un campo eléctrico las cargas positivas y negativas se mueven en direcciones opuestas y se genera un momento dipolar en cada molécula del material. Se puede definir como la polarización \mathbf{P} del material al momento dipolar neto promedio por unidad de volumen en el material, es decir, \mathbf{P} nos da una medida de cómo reacciona el material en conjunto al campo eléctrico externo [Sutherland, 2003]. En el caso de una onda, dado que el campo oscila, hace oscilar al momento dipolar. En la óptica lineal la relación entre el campo y la polarización se considera lineal como se muestra en la ecuación 2.1 para un campo constante. Ésto es una excelente aproximación en situaciones donde la intensidad es baja, como la luz que ocurre naturalmente. Éste es el régimen de la óptica lineal [Sutherland, 2003].

2.1. Polarización en el régimen lineal

$$\mathbf{P} = \epsilon_0 \chi^{(1)} \mathcal{E} \quad (2.1)$$

El símbolo $\chi^{(1)}$ en la ecuación 2.1 se conoce como el *tensor de susceptibilidad eléctrica*. Si queremos considerar a \mathcal{E} como una función en el tiempo (como es el caso si el campo externo es una onda oscilatoria) se debe considerar que el material no responderá de manera instantánea, y por lo tanto se debe considerar el *tensor de respuesta dieléctrica* $\chi^{(1)}(t - t')$, una función de $t - t'$ donde t es el tiempo donde se evalúa la polarización y t' es el tiempo en el cual se evalúa el campo externo^[1]. La dependencia de $\chi^{(1)}(t - t')$ respecto a esta variable es donde se considera el tiempo de respuesta, indica que a cierto tiempo t la respuesta del material al campo depende de cuánto tiempo antes se aplicó el campo (en t'). Para respetar la causalidad $\chi^{(1)}(t - t')$ debe ser cero cuando t' sea mayor (o igual) que t . Considerando esto, la polarización cuando se incide al material con una onda se puede escribir como se muestra en la ecuación 2.2 [Sutherland, 2003].

$$\tilde{\mathbf{P}}_L(\mathbf{r}, t) = \epsilon_0 \int_{-\infty}^{\infty} \chi^{(1)}(t - t') \cdot \mathcal{E}(\mathbf{r}, t') dt' \quad (2.2)$$

Donde la tilde encima de $\tilde{\mathbf{P}}_L(t)$ indica que oscila rápidamente. Si se supone que el tiempo de respuesta del material es muy rápido en comparación a la manera en la que varía la intensidad del campo, entonces se pueden ignorar las integrales, y simplemente usar algo como la ecuación 2.1 con dependencias del tiempo. Más adelante se supondrá esto por simplicidad. Resulta que (incluso sin suponer nada sobre el tiempo de respuesta) se puede escribir una relación más simple si se hace una transformada de Fourier y se escribe en el espacio de frecuencias^[2] [Sutherland, 2003]:

$$\bar{\mathbf{P}}_L(\omega) = \epsilon_0 \chi^{(1)}(\omega) \bar{\mathcal{E}}(\omega) . \quad (2.3)$$

^[1]En el caso más general se debe incluir también una dependencia de $\mathbf{r}-\mathbf{r}'$ que representa al material reaccionando de manera no local, es decir, reaccionando al campo en otra parte del material. Se supondrá que el material solo reacciona de manera local

^[2]las barras sobre P y E en la ecuación 2.3 indican que son las transformadas de Fourier

En la ecuación 2.3 la $\chi^{(1)}$ representa el tensor de susceptibilidad en función de la frecuencia, que es la transformada de Fourier del tensor de respuesta dieléctrica usado en la ecuación 2.2^[3].

2.2. Polarización en el régimen no lineal

Cuando la intensidad de la onda es suficientemente alta, llega un punto en el que las cargas ya no se pueden mover de manera proporcional al campo. Una de las maneras en las que puede ocurrir esto es que a cierta amplitud del campo las cargas en el material lleguen a un máximo de distancia de su punto de equilibrio y ya no puedan separarse más [Trebino, 2000]. Este caso es similar al “clipping” que ocurre cuando una señal pasa por un amplificador electrónico que no es capaz de alcanzar el voltaje suficiente para reproducir la señal. Se muestra este fenómeno en la figura 2.1. La señal original es aproximadamente “monocromática”, solo contiene una frecuencia. Por otro lado la señal distorsionada está compuesta de muchas frecuencias, contiene muchos armónicos de la frecuencia original [Trebino, 2000]. Así, cuando la polarización no puede seguir el comportamiento del campo externo linealmente, aparecen en $P(t)$ componentes de frecuencias que no se encuentran en el campo, y a su vez la polarización genera nuevas ondas electromagnéticas (después de todo cuando P oscila, eso significa que hay cargas aceleradas). Es así como, con suficiente intensidad, un material puede generar luz con frecuencias distintas a la frecuencia de la luz incidente cuando la respuesta del material ya no es lineal.

^[3]Cuando se evalúa este tensor de susceptibilidad en $\omega = 0$ obtenemos la ecuación en el caso constante, es decir, la ecuación 2.1

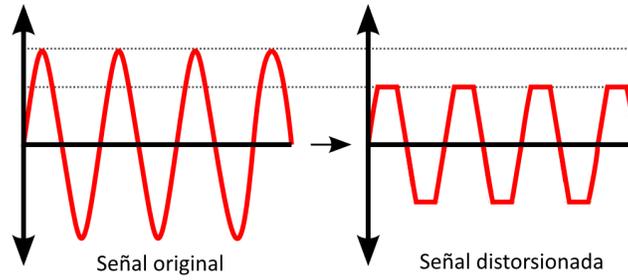


Figura 2.1: Representación de una señal distorsionada por “clipping”. La señal distorsionada no tiene forma senoidal, y por tanto contiene componentes de frecuencia distintas

En estos casos en los que la polarización no es una función lineal del campo como en la ecuación 2.1, se suele escribir como una serie de potencias, como la que se muestra en la ecuación 2.4 [Boyd, 2008].

$$\tilde{P}(t) = \epsilon_0 \left(\chi^{(1)} \mathcal{E}(t) + \chi^{(2)} \mathcal{E}(t)^2 + \chi^{(3)} \mathcal{E}(t)^3 + \chi^{(4)} \mathcal{E}(t)^4 + \dots \right) \quad (2.4)$$

En la ecuación 2.4 se ha escrito todo como escalares por simplicidad, esto es igual a suponer que el medio es isotrópico y por tanto, como ya se mencionó antes, $\chi^{(n)}$ son escalares, y los campos tienen la misma dirección y se puede tratar solo con sus amplitudes. Suele ser conveniente hablar de cada término por separado, y se usa la notación $P^{(n)}$ para hablar del n-ésimo término, y se le llama la polarización de orden n. Por ejemplo, la expresión para la polarización de orden 2, que se usará más adelante, es:

$$\tilde{P}^{(2)}(t) = \epsilon_0 \chi^{(2)} \mathcal{E}(t)^2. \quad (2.5)$$

2.3. SHG y suma de frecuencias

Veamos qué efectos ocurren cuando un material tiene susceptibilidad de orden 2. Si suponemos que el campo es aproximadamente monocromático (como lo sería un láser por ejemplo), podemos escribirlo de manera similar a como se hizo en el capítulo 1, en la

ecuación 1.4, pero sin omitir el complejo conjugado para asegurar que el campo sea real:

$$\mathcal{E}(t) = \frac{1}{2}E(t)e^{i\omega_0 t} + \frac{1}{2}E^*(t)e^{-i\omega_0 t} . \quad (2.6)$$

Si introducimos ese campo en la ecuación 2.5 se obtiene

$$\begin{aligned} \tilde{P}^{(2)}(t) &= \epsilon_0 \chi^{(2)} \left(\frac{1}{2}E(t)e^{i\omega_0 t} + \frac{1}{2}E^*(t)e^{-i\omega_0 t} \right)^2 , \\ \tilde{P}^{(2)}(t) &= \epsilon_0 \chi^{(2)} \left(\frac{1}{4}E(t)^2 e^{i2\omega_0 t} + \frac{1}{4}E^*(t)^2 e^{-i2\omega_0 t} + \frac{1}{2}E(t)E^*(t) \right) . \end{aligned} \quad (2.7)$$

Los primeros dos términos representan una onda con una frecuencia del doble de la frecuencia original. Esto muestra que ocurre lo que se conoce como *generación de segundo armónico* (o SHG, “second harmonic generation”) [Boyd, 2008]. Este fenómeno consiste, como lo indica su nombre, en la generación de un haz de luz con el doble de frecuencia del incidente. Este fue de hecho el fenómeno cuya observación en un cristal de cuarzo llevó al desarrollo de la óptica no lineal [Sutherland, 2003], y sigue siendo una herramienta particularmente importante para el campo de la óptica, en particular por sus usos en conversión de frecuencias y en la medición de pulsos [Diels and Rudolph, 2011], como se verá más adelante en este trabajo.

El tercer término de la ecuación 2.7 no tiene oscilaciones, es simplemente el producto de la amplitud compleja por su conjugado. Éste término corresponde a lo que se conoce como *rectificación óptica* [Boyd, 2008], ya que con la energía de la luz incidente se genera un campo constante en el tiempo. No es de mucho interés para este trabajo así que no se dirá más al respecto.

Si en vez de incidir luz monocromática incide al material incide luz con dos componentes monocromáticas, como podrían ser dos haces láser, entonces podemos escribir el campo en el punto donde incidan los haces como:

$$\mathcal{E}(t) = E_1(t)e^{i\omega_1 t} + E_2(t)e^{i\omega_2 t} + c.c. \quad (2.8)$$

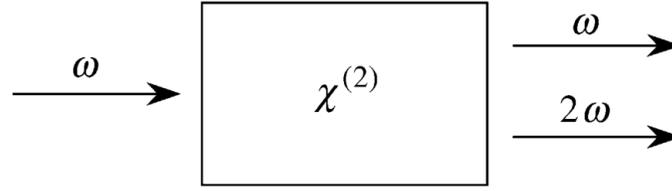


Figura 2.2: Diagrama representativo de la generación de segundo armónico. Entra al material un haz con frecuencia ω y salen dos haces, el transmitido con frecuencia ω y el segundo armónico, de frecuencia 2ω

Esta vez por simplicidad dejaremos indicados los complejos conjugados. Si introducimos este campo de 2.8 a la ecuación 2.5 obtenemos la ecuación 2.9 ^[4].

$$\begin{aligned}
 \tilde{P}^{(2)}(t) = & \epsilon_0 \chi^{(2)} \left(E_1(t)^2 e^{i2\omega_1 t} + E_2(t)^2 e^{i2\omega_2 t} + c.c. \right) \\
 & + \epsilon_0 \chi^{(2)} \left(2E_1(t)E_2(t)e^{i(\omega_1+\omega_2)t} + 2E_1(t)E_2^*(t)e^{i(\omega_1-\omega_2)t} + c.c. \right) \\
 & + 2\epsilon_0 \chi^{(2)} \left(E_1(t)E_1^*(t) + E_2(t)E_2^*(t) \right)
 \end{aligned} \tag{2.9}$$

En la ecuación 2.9 se pueden ver términos que corresponden a los fenómenos de los que ya se habló anteriormente, los primeros dos términos corresponden a SHG de cada componente y los últimos dos a la rectificación óptica de cada componente. Los términos restantes, que son términos cruzados que involucran a E_1 y E_2 , corresponden a los dos nuevos efectos que se observan en éste caso: *generación de suma de frecuencias* y *generación de diferencia de frecuencias*. Ambos efectos consisten en la generación de luz de nuevas frecuencias, $\omega_1 + \omega_2$ y $|\omega_1 - \omega_2|$ respectivamente, a partir de las dos frecuencias originales [Boyd, 2008].

Hasta ahora no hemos considerado ninguna característica espacial de la luz, sin embargo en el caso de dos haces aparece la posibilidad de que los haces no sean colineales, y debemos entonces considerar su dirección de propagación. Si no se suprime la parte

^[4]Se desarrollan completamente los cálculos involucrados en el anexo A

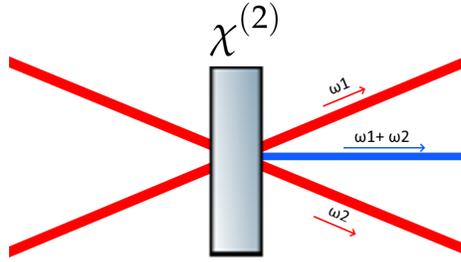


Figura 2.3: Diagrama que muestra los haces involucrados en el fenómeno de la suma de frecuencias.

espacial de la expresión de la onda, como se puede ver en el anexo A, los términos que corresponden a la generación de suma de frecuencias y a a generación de diferencia de frecuencias son:

$$\begin{aligned} \epsilon_0 \chi^{(2)} 2E_1(t)E_2(t)e^{i[(\omega_1+\omega_2)t - (\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{r}]} + c.c. \\ \epsilon_0 \chi^{(2)} 2E_1(t)E_2^*(t)e^{i[(\omega_1-\omega_2)t - (\mathbf{k}_1-\mathbf{k}_2)\cdot\mathbf{r}]} + c.c. \end{aligned} \quad (2.10)$$

Se puede ver en esas expresiones que los nuevos haces generados tendrán direcciones dadas por $(\mathbf{k}_1 + \mathbf{k}_2)$ y $(\mathbf{k}_1 - \mathbf{k}_2)$. En el caso de la suma de frecuencias en particular, el haz generado tendrá una dirección, mostrada con la línea color azul en la figura 2.3, coplanar con los dos haces y en una dirección intermedia entre los dos [Trebino, 2000].

Como nota para el resto de este trabajo, se usará en el transcurso de éste un arreglo en el cual dos haces con la misma frecuencia (que de hecho provienen de la misma fuente) se usan para generar un haz con el doble de frecuencia. El efecto que ocurre en este caso podría pensarse como SHG no-colineal o como una suma de frecuencias en la que ω_1 es igual a ω_2 . No es particularmente importante si se usa una categorización u otra, físicamente el efecto es completamente el mismo, sin embargo en este trabajo se considerará este caso como un ejemplo del efecto de suma de frecuencias. Nos referiremos al efecto como una suma de frecuencias simplemente para distinguirlo de la generación de segundo armónico que sucede en cada haz individualmente, y que resulta en haces colineales a ellos.

2.4. Phase matching

Al propagarse la luz incidente a través del material, induce una polarización a lo largo de él. Esta polarización, como se describió en la sección anterior, tiene asociada frecuencias y vectores de onda, por ejemplo el primero de los términos de la polarización de segundo orden mostrados en la ecuación 2.10 describe una onda con frecuencia $(\omega_1 + \omega_2)$ y vector de onda $(\mathbf{k}_1 + \mathbf{k}_2)$. La luz generada por esta polarización inducida debido al fenómeno de suma de frecuencias tendrá frecuencia $(\omega_1 + \omega_2)$, y la dirección de su vector de onda será la misma que la dirección de $(\mathbf{k}_1 + \mathbf{k}_2)$, sin embargo la magnitud de su vector de onda debe ser:

$$|\mathbf{k}_g| = \frac{(\omega_1 + \omega_2)n(\omega_1 + \omega_2)}{c}. \quad (2.11)$$

Donde c es la velocidad de la luz en el vacío, y $n(\omega_1 + \omega_2)$ es el índice de refracción del medio, evaluado en la frecuencia $(\omega_1 + \omega_2)$. Es claro entonces que, en general, la polarización inducida y la luz generada no tienen el mismo vector de onda, su magnitud es distinta, y por lo tanto no estarán en fase a lo largo del material. Cuando esto ocurre la luz generada en un punto del material no estará en fase con la luz que se genere en otro más adelante, y el efecto neto será que interfieran destructivamente, y por lo tanto no el material en conjunto no emitirá luz [Trebino, 2000]. Sólo se generará luz en fase a lo largo del material si la magnitud de \mathbf{k}_g coincide con la del vector de onda de la polarización inducida, y por lo tanto se empatan sus fases. A este concepto se le conoce como *empatamiento de fase*, o más comúnmente en inglés *phase matching*, y a la igualdad entre estos vectores de onda se le conoce como la condición de phase matching. El phase matching es una de las principales razones por la cual no se suelen observar efectos no lineales en la vida diaria; además de necesitarse un material con una respuesta no lineal, se necesitaría cumplir la condición de phase matching para que el efecto fuera visible, lo cual suele ser poco probable (si no imposible) a menos que se diseñe una situación específica para lograrlo.

2.4.1. Phase matching Para SHG

Para obtener la condición de phase matching, hay que conseguir que el vector de onda de la luz emitida sea igual al vector de onda del término correspondiente en la polarización inducida. Dado que el vector de onda de la polarización es una función de los vectores de onda de la luz incidente, y esa función es diferente para cada efecto, generalmente solo se puede establecer la condición de phase matching para un efecto a la vez. Consideremos ahora como ejemplo la generación de segundo armónico. El término correspondiente en la polarización, como se puede ver en el anexo A es:

$$\tilde{P}^{(2)}(\mathbf{r}, t) = \epsilon_0 \chi^{(2)} E_1^2 e^{i(2\omega_1 t - 2\mathbf{k}_1 \cdot \mathbf{r})} . \quad (2.12)$$

En ese caso la condición que se debe cumplir es:

$$2\mathbf{k}_1 = \mathbf{k}_s , \quad (2.13)$$

donde \mathbf{k}_1 es el vector de onda del haz incidente y \mathbf{k}_s es el vector de onda del haz de salida. Ésto nos dice que el haz de salida será colineal al haz incidente. Si se consideran las magnitudes de los vectores, y recordando que el haz de salida tiene una frecuencia del doble de la del haz incidente, entonces se tendrá la ecuación 2.14.

$$2 \frac{\omega_1 n(\omega_1)}{c} = \frac{2\omega_1 n(2\omega_1)}{c} , \quad (2.14)$$

y por tanto:

$$n(\omega_1) = n(2\omega_1) . \quad (2.15)$$

Aquí parece que se ha llegado a un callejón sin salida, la ecuación 2.15 indica que para poder observar SHG, se debe usar un material en el cual el índice de refracción sea igual para ω_1 y dos veces ω_1 . Es decir, parece que se necesita un material sin dispersión, donde el índice de refracción sea igual para toda frecuencia. Ésto es prácticamente imposible de conseguir, y ocurre lo mismo con el caso de la suma de frecuencias. Sin embargo resulta ser posible lograr la condición en el caso de un medio birrefringente. Los medios birrefringentes tienen un índice de refracción diferente dependiendo de la polarización de la luz que

los atraviesa, así que si se polariza linealmente el haz incidente se puede lograr que un haz de salida con la polarización ortogonal, y que por tanto ve un índice de refracción distinto, cumpla la ecuación 2.15. Para lograr esta condición se debe usar el material birrefringente en una orientación específica que depende de la frecuencia a la que se quiere lograr la condición de phase matching. Los fabricantes de elementos ópticos hechos específicamente para obtener efectos no lineales normalmente cortan un cristal birrefringente de manera que sea sencillo orientarlos para obtener el phase matching con cierta frecuencia.

Para casos como la suma de frecuencias o un arreglo para obtener SHG no-colineal ^[5] existen dos maneras de lograr que se cumpla la relación correspondiente usando un material birrefringente, y se conocen como phase matching de *tipo 1* y *tipo 2*. En el phase matching tipo 1 los dos haces incidentes se polarizan en la misma dirección, y se orienta el cristal de manera que el haz de salida cumpla la relación cuando tenga la polarización ortogonal. En el phase matching tipo 2 se modifica la polarización de uno de los haces incidentes, y el haz de salida tendrá la misma polarización que alguno de ellos. La elección de un tipo sobre otro es meramente logística, pero dependiendo de los materiales disponibles y de la frecuencia que se quiere obtener puede ser una elección sumamente importante.

Otra consideración importante es que la condición de phase matching solo se cumple en realidad para una frecuencia específica. Sin embargo si se recuerda la discusión del inicio de esta sección, cuando no se da la condición de phase matching lo que ocurre es que la polarización se empieza a desfazar de la luz generada a lo largo de su propagación. Si se analiza una frecuencia cercana a la que cumple la condición de phase matching se observaría que el desfase ocurre lentamente, y si el medio en el que ocurre el efecto es suficientemente delgado, el desfase nunca llega a ser suficiente para interferir destructivamente y evitar que la luz sea emitida al final del medio. Esto significa que existe efectivamente un ancho de banda de phase matching que depende del grosor del medio que se use, entre más delgado el medio usado, mayor ancho de banda se tendrá. Por otro lado, es claro que entre más grueso sea el medio, el efecto no lineal en general se verá más fuertemente, pues hay más material que polarizar y que por tanto se emitirá más luz. Esto significa que hay que considerar cuidadosamente el grosor del medio usado en la práctica.

^[5]En el caso colineal solo hay un haz de entrada y por lo tanto solo existe la opción de modificar la polarización del haz de salida respecto a éste

Para estimar el ancho de banda de phase matching para un grosor dado se encuentra que, como primera aproximación, para el caso donde la potencia de la luz generada es mucho menor que la potencia de la luz incidente, la eficiencia de la conversión como función de la diferencia entre los vectores de onda se comporta como[Sutherland, 2003]:

$$\frac{\sin^2(\Delta k L / 2)}{(\Delta k L / 2)^2} . \quad (2.16)$$

Se considera como el ancho de banda de phase matching al intervalo donde esta cantidad tiene valor mayor a $\frac{1}{2}$. Esto ocurre aproximadamente en el intervalo donde $-2.78 \leq \Delta k L \leq 2.78$.

Para el caso de la generación de segundo armónico, la diferencia Δk se puede escribir, considerando la ecuación 2.14 como:

$$\Delta k = 2 \frac{\omega_1 n(\omega_1)}{c} - \frac{2\omega_1 n(2\omega_1)}{c} . \quad (2.17)$$

Convirtiendo a longitud de onda y simplificando se obtiene:

$$\Delta k = \frac{4\pi}{\lambda} [n(\lambda) - n(\lambda/2)] . \quad (2.18)$$

Si se considera un cambio $\delta\lambda$ al rededor de la longitud de onda donde ocurre el phase matching, λ_0 , y se calcula el intervalo tal que Δk caiga en los valores antes mencionados se encuentra que el ancho de banda de phase matching, en términos de la longitud de onda y en función del ancho del material L, es[Trebino, 2000]:

$$\delta\lambda_{pm} = \frac{0.44\lambda_0/L}{|n'(\lambda_0) - \frac{1}{2}n'(\lambda_0/2)|} . \quad (2.19)$$

Donde $n'(\lambda_0)$ es la derivada del índice de refracción respecto a la longitud de onda, evaluada en la longitud de onda λ_0 . Si se trata de phase matching tipo 1, es importante

notar que la función $n(\lambda)$ no es la misma para la luz incidente y la generada, debido a su polarización y a la birrefringencia del medio. No resulta sencillo obtener los valores de n' para un caso particular, pues el índice de refracción del haz extraordinario dependerá del ángulo respecto al eje óptico.

Considérese el caso de SHG con un cristal de beta borato de bario (BBO), que como se mencionará más adelante es el material que se usó en este trabajo. Se pueden encontrar en la literatura [Eimerl et al., 1987] valores de n_o , n_e , y sus respectivas derivadas para el BBO. Como aproximación se usará n'_o para el haz incidente, que es el ordinario, y n'_e para el haz generado^[6]. Un pulso $sech^2$ ^[7] de 800nm limitado en ancho de banda con duración de 100fs tiene un ancho de banda de aproximadamente 10 nm. Así, despejando L de la ecuación 2.19 se puede obtener una aproximación al grosor máximo que se debería usar para un caso similar al de este trabajo de:

$$L = \frac{0.44(800nm/10nm)}{|(-0.030\mu m^{-1}) - \frac{1}{2}(-0.158\mu m^{-1})|} \approx 720\mu m . \quad (2.20)$$

Los fabricantes recomiendan que para pulsos como el que se describió en el parrafo anterior, de duración de 100fs a 800nm, se use un cristal BBO con un grosor de 500 micras [EKSMA-Optics, sf].

^[6]En realidad el índice que verá el haz generado es un valor entre n_o y n_e

^[7]Un pulso cuya intensidad como función del tiempo se aproxima a la función secante hiperólica al cuadrado

Capítulo 3

La técnica FROG

En general, cuando se quiere medir la duración de un evento o un proceso, se necesita tener un evento de menor duración con el cual compararlo, así como para medir la longitud de un objeto se necesita compararlo con las marcas en una regla. Los pulsos ultracortos de láser, de los que se habló en el capítulo 1, pueden llegar a ser los eventos de duración más corta que se han podido generar en un laboratorio. Por dar algunos ejemplos, se han llegado a generar pulsos directamente a la salida de un láser con duraciones de hasta 5fs [Kärtner et al., 2004], y con la adición de complejos arreglos experimentales se pueden generar pulsos hasta del orden de cientos de attosegundos [Schultze et al., 2007]. Por lo tanto surge la pregunta: ¿cómo se puede medir la duración de uno de estos pulsos? La respuesta a esta pregunta por mucho tiempo, antes de la creación de la técnica FROG, fue que solo se puede estimar usando la *autocorrelación no lineal*. Para entender la técnica usada en este trabajo, conocida en inglés como *frequency resolved optical gating*^[1] (FROG), es necesario comenzar entendiendo qué es la autocorrelación no lineal (a la cual desde ahora se le referirá simplemente como autocorrelación), pues FROG fue creada a partir de ella.

Si se tuviera un evento de duración mucho menor al pulso que se quiere medir, sería relativamente sencillo medir la intensidad y duración del pulso. Bastaría con usar el evento

^[1]Resulta un poco complicado traducir este término de manera fácil, muchos de los autores de habla hispana suelen no hacerlo, pero se podría traducir como “gating óptico resuelto en frecuencia” o “método de compuerta óptica resuelto en frecuencia”

más corto como “compuerta” ^[2] de manera que con un sensor se midiera la intensidad (o más probablemente la potencia) del pulso sólo durante la duración de ese evento más corto. Si se traslapara esta compuerta con el pulso en diferentes partes del pulso, es decir si se retarda uno respecto al otro de manera variable, se podría obtener una medida de la intensidad del pulso respecto al tiempo, justo lo que se busca. La autocorrelación usa esta idea, y resuelve el problema de no tener un evento más corto usando una réplica del pulso mismo como la compuerta.

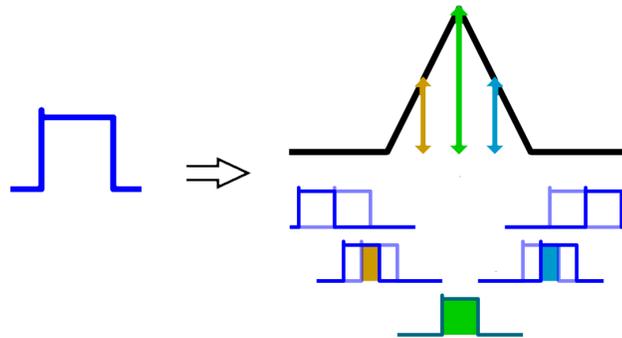


Figura 3.1: Diagrama de la autocorrelación de una señal.

Se muestra en la figura 3.1 un diagrama de la autocorrelación de una señal. La señal se duplica y se recorre respecto a la original. Donde las señales no se empalman, la autocorrelación es cero. Donde se empalman la autocorrelación aumenta, y su máximo se encuentra donde las señales están completamente empalmadas.

Para poder usar un pulso de luz como compuerta se debe usar un medio óptico no lineal (generalmente de segundo orden). Se necesita un medio no lineal porque como ya se habló en el capítulo 2, en la óptica no lineal existen maneras de que un haz de luz interactúe con otro (o consigo mismo) mediante un medio, esto no es posible en la óptica lineal. Hay muchos efectos ópticos no lineales que se pueden usar para este fin, pero en este trabajo se considerará sólo la generación de segundo armónico (SHG).

^[2]Gate en inglés

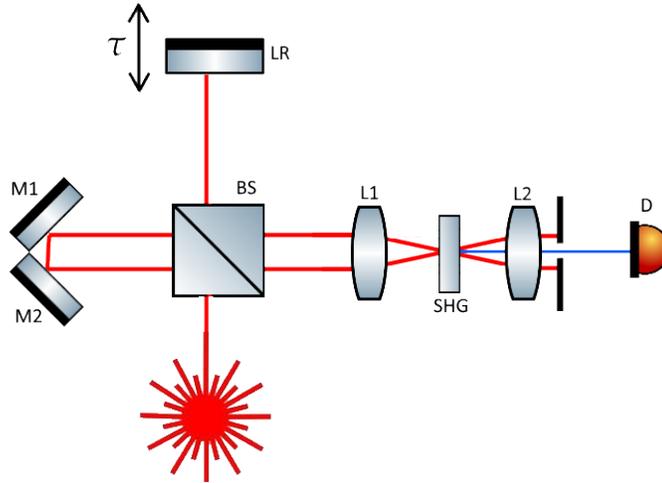


Figura 3.2: Diagrama de un autocorrelador de intensidad por SHG.

Para medir la autocorrelación de intensidad de un pulso, denotada^[3] $A^{(2)}(\tau)$, se usa un arreglo como el que se muestra en la figura 3.2. Primero se divide un pulso usando un divisor de haz (BS), se pasa uno de los haces por un camino de longitud variable (hacia la línea de retardo LR) y el otro se desvía con dos espejos (M1 y M2)^[4]. Posteriormente los dos haces se enfocan (con la lente L1) en un medio que pueda generar su suma de frecuencias, en este caso un medio que presenta generación de segundo armónico (SHG)^[5] a la frecuencia del haz, pues ambos pulsos tienen la misma frecuencia. Después del medio se colima el haz resultante (con L2), se tapan los haces fundamentales con un diafragma y se mide la suma de frecuencias (con el detector D) variando la diferencia de longitud del camino óptico, y por tanto el retardo relativo entre los pulsos τ . Si el campo del pulso es $E(t)$, la amplitud del campo generado por la suma de frecuencias cumple:

$$E_{SFG} \propto E(t)E(t - \tau), \quad (3.1)$$

^[3]Donde el superíndice (2) indica que se trata de una autocorrelación de segundo orden

^[4]Se podría mostrar aquí un solo espejo con un ángulo de 90 grados, sin embargo en la práctica suele ser útil que el ángulo entre M1 y M2 se pueda variar ligeramente para corregir desviaciones en la alineación de otras partes del arreglo

^[5]Como se mencionó en la página 15 se ha llamado al efecto que se usará suma de frecuencias, pues se usan dos haces, cada uno de los cuales también podrá presentar su propia SHG. Al describir al cristal, lo que se busca es que presente SHG a la frecuencia relevante, y naturalmente también podrá generar la suma de frecuencias en este caso, pues es una suma con dos frecuencias iguales.

y por lo tanto la intensidad cumple que:

$$I_{SFG} \propto I(t)I(t - \tau) . \quad (3.2)$$

Dado que ningún medidor es tan rápido como para discernir la dependencia respecto al tiempo, lo que se medirá será:

$$A^{(2)}(\tau) = \int_{-\infty}^{\infty} I(t)I(t - \tau) dt . \quad (3.3)$$

La ecuación 3.3 es la definición de la autocorrelación. Es claro que la autocorrelación puede dar una idea del ancho temporal del pulso, pues cuando se aumenta τ hasta que ya no se enciman los pulsos $A^{(2)}(\tau)$ será cero. Sin embargo no es claro cómo se relaciona la forma del pulso, es decir su intensidad respecto al tiempo, con la forma de $A^{(2)}(\tau)$. De hecho resulta que hay varias formas de pulso que resultan en la misma autocorrelación, entonces es imposible obtener en general $I(t)$ del pulso a partir de $A^{(2)}(\tau)$.

Descrito de la manera más sencilla, una medida FROG es simplemente una autocorrelación donde adicionalmente se resuelve la señal en frecuencias, es decir, se pasa la luz resultante de la autocorrelación por un espectrómetro. Dado que en este caso se está usando como no linealidad la generación de segundo armónico en el autocorrelador, la técnica se conoce como SHG-FROG. Por lo tanto, si la intensidad de la autocorrelación era dada por la ecuación 3.2, en SHG-FROG se obtiene el espectro para cada τ , es decir, se hace una transformada de Fourier a 3.1 respecto a t , se eleva al cuadrado para obtener la intensidad, y se obtiene la ecuación 3.4.

$$I_{FROG}(\omega, \tau) = \left| \int_{-\infty}^{\infty} E(t)E(t - \tau) \exp(-i\omega t) dt \right|^2 \quad (3.4)$$

El resultado es función del retardo τ y de la frecuencia ω , y por lo tanto se debe representar en una imagen 2D, conocida como *la traza FROG*. Resulta ser que la traza FROG contiene suficiente información extra comparada con la autocorrelación como para poder

reconstruir completamente $I(t)$ del pulso [Trebing, 2000]. Se discutirá esto con más detalle en el siguiente capítulo.

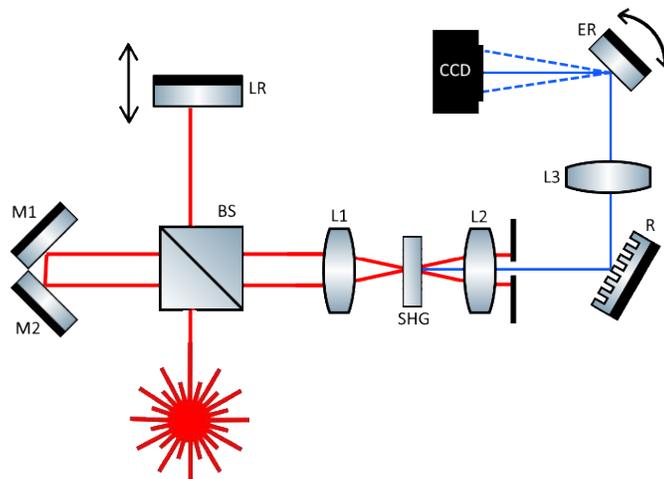


Figura 3.3: Diagrama de un arreglo SHG-FROG. Es igual al arreglo de la autocorrelación excepto que se agrega una sección de espectrómetro, que consiste en la rejilla R, la lente L3, y el espejo rotatorio ER. El resultado se capta en una CCD.

Se muestra en la figura 3.3 un diagrama de un arreglo FROG que usa un cristal que presenta SHG como medio no lineal. De igual manera que en la autocorrelación, se divide el haz en dos. Uno de los haces pasa por un camino cuya longitud es variable, a esta sección se le llama la *línea de retardo*. Posteriormente se enfocan sobre el cristal no lineal. Cuando en ambos haces hay un pulso pasando por el cristal simultáneamente, se generará la suma de frecuencias. Este nuevo haz se aísla de los haces fundamentales con un diafragma y se manda hacia una rejilla de difracción. La luz posteriormente pasa por la lente L3 y se redirige con un espejo rotatorio hacia una cámara CCD, ubicada en el plano de Fourier, es decir, a una distancia de L3 igual a su distancia focal. La distancia entre la rejilla y la lente L3 determinará el tamaño que el espectro tendrá en la cámara, y por tanto, como se verá mas tarde, afectará la calibración del dispositivo^[6]. El espejo rota en sincronía con el movimiento de la línea de retardo, de manera que para cada posición de la línea de

^[6]Dado que la luz fue colimada por la lente L2, la distancia entre la rejilla de difracción y L3 no afectará la distancia a la que la imagen del espectro se enfoca, sólo su tamaño. Cabe mencionar también que no se necesita de una rendija a la "entrada" del espectrómetro, pues la fuente de la luz que se dirige al espectrómetro, el punto de incidencia de los haces láser en el cristal no lineal, es aproximadamente puntual, y es la imagen de éste que la lente L3 enfoca en el plano de Fourier (claro, con cada componente espectral recorrida por la acción de la rejilla)

retardo, cada τ , la luz incide en una posición diferente de la cámara, formando de los espectros correspondientes a cada retardo la traza FROG. [Garduño-Mejía et al., 2003]

3.1. La traza FROG

El propósito del espejo rotatorio que se mencionó anteriormente es redirigir el espectro obtenido por la rejilla hacia diferentes partes de la cámara conforme se varía la línea de retardo, de manera que resulte una imagen 2D que contenga en un eje información temporal, correspondiente al retardo relativo entre los pulsos y a la rotación del espejo, y en el otro eje información relativa a la frecuencia. La traza FROG entonces es esencialmente una gráfica de intensidad contra dos parámetros, retardo τ y frecuencia.

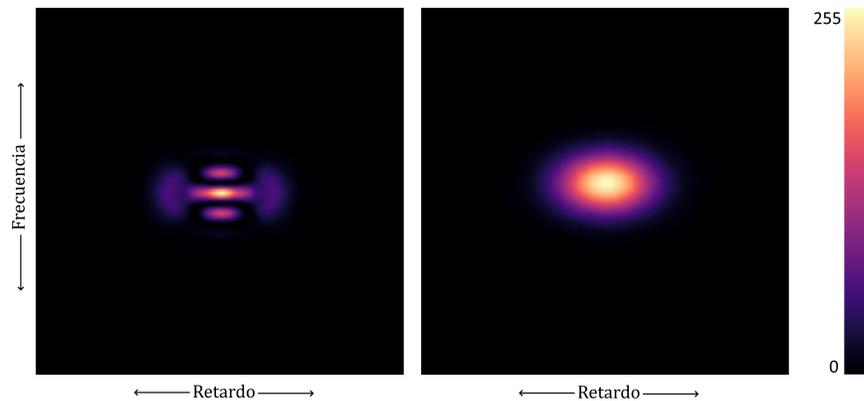


Figura 3.4: Dos ejemplos de trazas SHG-FROG. Se usó color falso para indicar la intensidad, a la derecha se muestra el mapa de color usado. El eje horizontal representa el retardo y el eje vertical representa la frecuencia.

En la figura 3.4 se muestran dos ejemplos de trazas FROG simuladas ^[7]. El eje horizontal representa el retardo, y el eje vertical la frecuencia, de manera que cortes verticales de la imagen representan el espectro de la luz que sale del cristal para cada retardo relativo dado. Se usó un mapeo de color para hacer mas claras las diferencias sutiles de intensidad. Estas trazas son simuladas, y por lo tanto sus ejes no están marcados ni tienen unidades. En el caso de una traza real sería posible indicar la escala una vez calibrado el arreglo.

^[7]Para más información sobre cómo se simularon las trazas, refiérase al capítulo 5

Las Trazas SHG-FROG apropiadas para ser analizadas deben cumplir las siguientes características:

- Están completamente centradas en la imagen
- Son simétricas en el eje horizontal (Este hecho está relacionado con la ambigüedad en la dirección del tiempo que se mencionó anteriormente)
- Se vuelven cero en las orillas

Una traza obtenida mediante la técnica SHG-FROG que no cumpla estas características no corresponde a una traza real, o no contiene suficiente información para asociarla con un pulso real. Esto significaría que hay errores sistemáticos en la medida.

Una manera de diagnosticar errores sistemáticos y contrarrestarlos es usando los *marginales* de la traza FROG, que son las integrales de la traza sobre uno de sus ejes. Al calcular los marginales se obtienen cantidades relacionadas al espectro (si se integra sobre el eje horizontal, en cuyo caso se conoce como el marginal de frecuencia) y a la autocorrelación de intensidad (si se integra sobre el eje vertical, en cuyo caso se conoce como el marginal de retardo) [Trebino, 2000]. En particular para el FROG SHG, el marginal de frecuencia debe ser igual a la autoconvolución del espectro, es decir [Trebino, 2000]:

$$M_{\omega}(\omega - 2\omega_0) = S(\omega - \omega_0) * S(\omega - \omega_0) = \int_{-\infty}^{\infty} S(\Omega)S(\omega - 2\omega_0 - \Omega) d\Omega . \quad (3.5)$$

Mientras que es fácil ver por la construcción de la traza que el marginal de retardo debe ser simplemente la autocorrelación del pulso. Es importante corroborar que se cumplen estas relaciones al menos una vez para asegurar que el aparato usado no tiene errores sistemáticos. Además en el caso de SHG-FROG si el marginal de frecuencia no coincide con la autoconvolución del espectro, se puede corregir la presencia de los errores sistemáticos que causaron que no coincida simplemente multiplicando la traza por el cociente de la autoconvolución y el marginal para cada frecuencia, forzando estos valores a coincidir. De

esta manera se pueden usar las trazas obtenidas aunque el espectrómetro usado o el cristal no lineal no tengan una respuesta uniforme en frecuencia.

Capítulo 4

El algoritmo FROG

Ya que hemos entendido lo que representa la traza y cómo está relacionada al pulso, es el momento de obtener la información del pulso a partir de ella. Para obtener la amplitud y fase del pulso a partir de la traza se usa un algoritmo iterativo de análisis, conocido como el algoritmo FROG. Existen muchas versiones del algoritmo, que con el tiempo han sido modificadas y mejoradas en términos de eficiencia y velocidad. Para entender cómo y porqué funcionan lo mejor es comenzar explicando la forma más básica del algoritmo, la que fue usada desde la creación de la técnica.

4.1. El algoritmo FROG básico

Considérese la Ecuación 3.4, y llamemos función señal a lo que está dentro de la integral, exceptuando la exponencial.

$$E_{sig}(t, \tau) = E(t)E(t - \tau) \quad (4.1)$$

Si se piensa a la función señal como la transformada de Fourier (sólo respecto a Ω , asociada con τ) de una función $\bar{E}_{sig}(t, \Omega)$, entonces la traza FROG está dada por la Ecuación

4.2

$$I_{FROG}(\omega, \tau) = \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \bar{E}_{sig}(t, \Omega) \exp(-i\omega t - i\Omega\tau) dt d\Omega \right|^2. \quad (4.2)$$

De esta manera, la traza FROG se puede entender como la transformada de Fourier en dos dimensiones de la función $\bar{E}_{sig}(t, \Omega)$. Si se tiene la traza y se pudiera obtener a partir de ella la función $\bar{E}_{sig}(t, \Omega)$, sería relativamente sencillo conocer completamente el pulso $E(t)$; si se le aplica una transformada de Fourier a $\bar{E}_{sig}(t, \Omega)$ respecto a Ω , y se evalúa en $\tau = t$, el resultado sería el pulso multiplicado por un factor constante complejo $E(0)$. En este caso, un factor constante no es de interés, pues no se busca conocer la intensidad absoluta ni la fase absoluta del pulso, sino sus variaciones, y por lo tanto conocer $\bar{E}_{sig}(t, \Omega)$ es equivalente a conocer el pulso original. Así, vemos que para conocer el pulso original a partir de la traza FROG se debe resolver un problema equivalente al de obtener una función de dos variables a partir sólo de la intensidad de su transformada de Fourier en dos dimensiones. A este problema se le conoce como *recuperación de fase en dos dimensiones* (o *two dimensional phase retrieval* en inglés) [Trebino, 2000].

El problema de la recuperación de fase en dos dimensiones aparece en otras áreas donde puede ser necesario reconstruir una imagen a partir de información relacionada a su espectro, por ejemplo en cristalografía, microscopía de electrones y astronomía [Fienup, 1982], y de hecho es un problema resuelto dada cierta información adicional sobre la función a recuperar [Trebino, 2000]. En el caso del FROG por SHG la información adicional que se tiene es que la función señal tiene la forma dada en la Ecuación 4.1. Esto resulta ser suficiente información para que haya poca ambigüedad sobre la forma del pulso dada una traza FROG. Aunque aún no existe una prueba matemática de la unicidad de la solución del FROG, en general se considera que el algoritmo recupera una única solución para una traza dada salvo tres ambigüedades “triviales”:

- Un factor constante de fase, i.e. $e^{i\phi_0} E(t)$
- Una traslación en el tiempo, i.e. $E(t + t_0)$

- Una inversión en la dirección del tiempo, i.e. $E^*(-t)$ [Trebino, 2000]

Los algoritmos que se usan más comúnmente para analizar trazas FROG suelen tener una estructura como la que se muestra en la figura 4.1. El algoritmo es iterativo, comienza con una propuesta arbitraria del pulso $E(t)$, y se itera un proceso que hace que cada vez el pulso propuesto se aproxime más al pulso medido. [Trebino, 2000] [Kane, 2008] De particular importancia son los pasos que mencionan la *restricción de intensidad* y la *restricción de forma matemática*. Es en el uso de estas restricciones que se consideran los datos obtenidos de la medición como tal, es decir, de la traza FROG experimental, y el conocimiento teórico de cómo se construye la traza a partir del pulso, respectivamente.

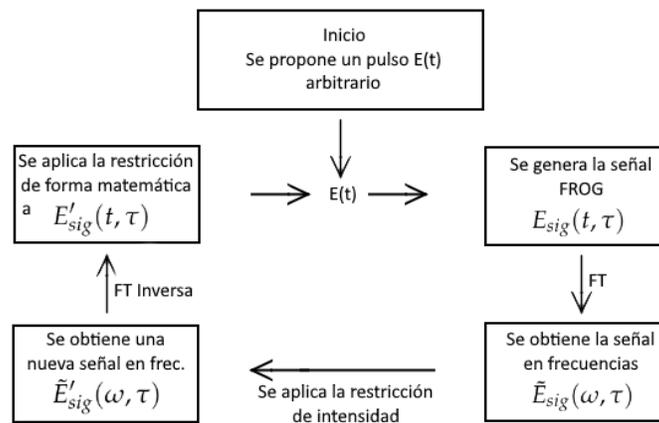


Figura 4.1: Diagrama de la estructura general de un algoritmo FROG genérico

Dado que la manera de medición de cualquier cantidad involucrada requiere de discretizar t y τ , consideremos que el pulso propuesto es un vector $\mathbf{E} = [E_1, E_2, E_3, \dots, E_N]$, donde cada componente representa la amplitud compleja del pulso a un tiempo Δt después de la anterior. Para generar la señal FROG $E_{sig}(t, \tau)$, dada en la ecuación 4.1 para el caso de FROG por SHG, se calcula el producto exterior de este vector \mathbf{E} consigo mismo. El resultado es una matriz de $N \times N$ que contiene la misma información que $E_{sig}(t, \tau)$. Posteriormente se realiza una transformación que consiste en “rotar” el i -ésimo renglón i lugares de tal manera que el elemento $E_i E_i$ sea el primer elemento en cada renglón, y reacomodar las columnas de tal manera que vayan en orden creciente respecto al valor de τ que repre-

sentan. Una vez realizada ésta transformación, la matriz resultante es la versión discreta de $E_{sig}(t, \tau)$, es decir, la representación en el espacio temporal del espectrograma formado por el aparato SHG FROG para el pulso $E(t)$. El siguiente paso es obtener $\tilde{E}_{sig}(\omega, \tau)$ haciendo la transformada de Fourier discreta respecto a t de la matriz. [Kane, 2008]

La restricción de intensidad simplemente requiere comparar la traza obtenida mediante la medición con la traza obtenida del pulso arbitrario propuesto. La manera más común de hacer esto es simplemente reemplazando la magnitud de cada elemento de la matriz $\tilde{E}_{sig}(\omega, \tau)$ con la raíz cuadrada de la intensidad del elemento correspondiente de la traza FROG medida. [Trebino, 2000] Para hacer esto es importante claramente que la matriz tenga el mismo tamaño que la traza medida, así que se debe escoger cuidadosamente el tamaño del vector \mathbf{E} tomando en cuenta los datos medidos. Una vez hecho esto se hace la transformada de Fourier inversa, obteniendo así una matriz señal nueva $E'_{sig}(t, \tau)$.

La manera de obtener una nueva propuesta de pulso para la siguiente iteración a partir de $E'_{sig}(t, \tau)$ es donde los algoritmos comunes difieren significativamente. Esto es la aplicación de la restricción de forma matemática, y el cómo se aplica depende de qué algoritmo se esté usando, pero la meta siempre es forzar a la matriz señal a que tenga la forma dada por la ecuación 4.1 o encontrar un pulso que, mediante dicha ecuación, se aproxime lo más posible a la matriz. Es decir, en este paso lo que se hace es restringir a la matriz señal a que tenga sentido físico según lo que se sabe de como un pulso generaría esa señal FROG, y recuperar ese pulso. En el primer algoritmo que se usó con la técnica FROG [Trebino and Kane, 1993], este paso consistía simplemente en integrar la señal respecto a τ , de manera que en cada iteración k :

$$E'(t) = \int_{-\infty}^{\infty} E'_{sig}(t, \tau) d\tau . \quad (4.3)$$

Notando que si la señal realmente tiene la forma dada por la ecuación 4.1, el resultado será simplemente un múltiplo escalar del pulso $E(t)$ [Trebino, 2000]. Así, usando esta nueva propuesta de pulso, se repite el proceso desde el inicio.

Para cuantificar el progreso en cada iteración, se puede definir una cantidad que mide el "error" entre las trazas frog generadas por el pulso real y el pulso recuperado por el

algoritmo. Esta cantidad, conocida como el *error FROG*, se define en la iteración k como:

$$G^{(k)} = \sqrt{\frac{1}{N^2} \sum_{i,j=1}^N \left| I_{FROG}(\omega_i, \tau_j) - \mu |\tilde{E}_{sig}^{(k)}(\omega_i, \tau_j)|^2 \right|^2}. \quad (4.4)$$

Donde μ es una constante real de normalización que minimiza $G^{(k)}$, $\tilde{E}_{sig}^{(k)}(\omega, \tau)$ es la matriz señal en el espacio de frecuencias en la k -ésima iteración, y N es su tamaño en ambas dimensiones (i.e. es una matriz de $N \times N$) [Trebino, 2000].

4.2. El algoritmo FROG por PCGP

La principal diferencia entre la mayoría de los algoritmos que se usan con frecuencia es en cómo se aplica la restricción de forma matemática. Como ya se mencionó, lo que se busca al aplicar esta restricción en cada iteración es obtener el pulso hipotético que se usará como propuesta en la siguiente iteración del proceso. En el método de *proyecciones generalizadas* [DeLong et al., 1994], se busca aplicar esta restricción mediante proyecciones. Si imaginamos el conjunto geométrico de todas las posibles funciones complejas que cumplen la restricción, i.e. que cumplen la ecuación 4.1, dentro del espacio de todas las funciones complejas, una proyección generalizada es una forma de modificar nuestra propuesta de pulso que dé como resultado la función más cercana que se encuentra dentro del conjunto. De igual manera podemos pensar en el conjunto de todas las funciones complejas que cumplen la restricción de intensidad y aplicarla como una proyección. Si la solución es única, entonces los dos conjuntos comparten sólo un punto, correspondiente a una función compleja que describe un pulso cuya traza FROG es igual a la traza medida, es decir, un pulso igual al pulso medido. El proceso de proyecciones sucesivas se acercará al punto de intersección como se muestra esquemáticamente en la figura 4.2. Se puede demostrar que si los dos conjuntos son convexos, entonces un método de proyecciones siempre convergerá a la solución [Trebino, 2000].

Sin embargo, en el caso del FROG, estos dos conjuntos en general no son convexos, y por lo tanto la proyección no necesariamente está bien definida, por lo cual se debe recurrir

a definir una proyección generalizada[Trebino, 2000], de ahí el nombre del método.

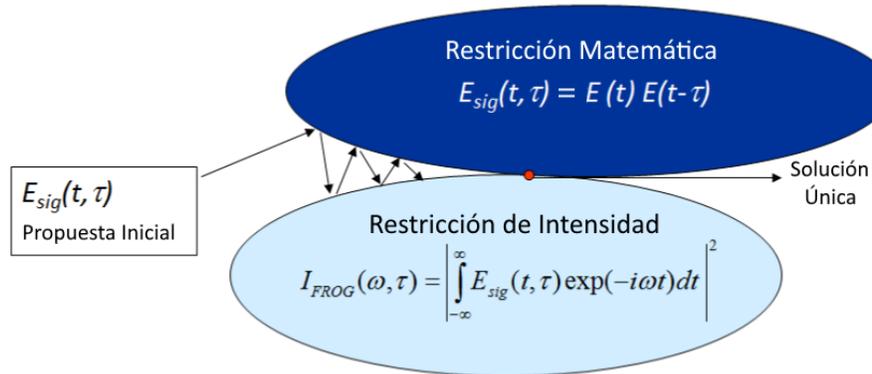


Figura 4.2: Diagrama del concepto detrás del funcionamiento de un algoritmo de proyecciones generalizadas

El método de proyecciones generalizadas tiene la desventaja de que, para definir la proyección, se debe hacer una minimización en un espacio de N dimensiones, donde N es el número de coordenadas del vector \mathbf{E} que se definió en la sección anterior. N está ligada a la resolución de la imagen de la traza que se introducirá al algoritmo, por dar un ejemplo en el caso de este proyecto $N = 256$. Es claro que minimizar una distancia en un espacio de 256 dimensiones es un cálculo computacionalmente pesado, lo cual reduce drásticamente la velocidad de este método, y adicionalmente hace que la programación sea complicada. El *algoritmo de proyecciones generalizadas por componentes principales* (algoritmo PCGP) busca remediar esta desventaja, siendo un método que se beneficia de las proyecciones generalizadas sin necesidad de hacer una minimización directamente [Kane, 2008].

En la sección anterior se dio por sentado que los dos pulsos que se juntan en el cristal son iguales, sin embargo el algoritmo PCGP no presupone este hecho, por lo cual funciona también para casos donde no lo son, como lo es la técnica TREEFROG [Kane, 2008], donde se usan dos pulsos distintos, uno como señal y otro independiente como compuerta, y se recuperan ambos. Para describir el algoritmo, por tanto, es más claro pensar en el caso en el que son independientes. Se les llamará $E(t)$ y $G(t)$ a los pulsos en cuestión, y más tarde se tomará en cuenta su relación.

El algoritmo tiene la misma estructura que el algoritmo descrito en la sección 4.1, solo difiere en el último paso, y en el hecho de que se consideran el pulso señal \mathbf{E} y el pulso compuerta \mathbf{G} independientes. Para obtener la matriz señal de estos vectores, se sigue el mismo proceso descrito en la sección ya mencionada, salvo que el producto exterior se hace entre \mathbf{E} y \mathbf{G} . Este procedimiento para encontrar la matriz señal es crucial para el funcionamiento del algoritmo, en particular es importante notar que la matriz señal se obtiene solamente haciendo operaciones simples al producto exterior que no cambian sus eigenvectores y eigenvalores.

El procedimiento continúa igual, se transforma la matriz señal al espacio de frecuencias, se modifica usando la traza medida y transformando de regreso al espacio de tiempo se obtiene una nueva matriz señal. Recordemos que lo que se busca es obtener una nueva propuesta de \mathbf{E} y \mathbf{G} a partir de esta nueva matriz señal. Como ya se mencionó, el proceso de obtener la matriz señal a partir de los pulsos no cambia los eigenvectores y eigenvalores, y además es un procedimiento invertible, es decir, es en principio posible obtener \mathbf{E} y \mathbf{G} a partir de la matriz señal. Para hacer esto, solo se necesita recordar que una matriz obtenida por el producto exterior de dos vectores tiene rango uno, es decir, solo tiene un eigenvalor distinto de cero, y el eigenvector asociado es uno de los vectores. El vector restante es el eigenvector izquierdo del producto exterior, i.e. el complejo conjugado del eigenvector de la matriz transpuesta [Friedberg et al., 1982]. Por tanto, en principio basta con obtener los eigenvectores de la matriz señal y su transpuesta. El caso de nuestro algoritmo resulta no ser así de sencillo, eso sería cierto para la matriz señal real, sin embargo la matriz obtenida en cualquier iteración dada al aplicar la restricción de intensidad no es una matriz de rango uno que se obtiene del producto exterior de dos vectores. Si lo fuera, eso significaría que es la matriz señal real, es decir, que ya fue resuelto el problema. En una iteración cualquiera la matriz señal obtenida no es de rango uno, lo más probable es que tenga N eigenvectores asociados, entonces se debe encontrar una manera de obtener vectores \mathbf{E} y \mathbf{G} apropiados a partir de ella.

Para entender cómo se logra esto, supongamos que se tiene una matriz \mathbf{O} , y que la

escribimos como un producto de tres matrices:

$$O = UWV^T, \quad (4.5)$$

donde U y V son matrices ortogonales cuadradas y W es una matriz diagonal (también cuadrada). Esto se conoce como *descomposición en valores singulares* (SVD), es equivalente a descomponer O en una serie de productos exteriores entre las columnas de U y los renglones de V , donde los elementos de la diagonal de W son factores que multiplican a cada producto externo. El producto externo que corresponde al valor más grande de W (i.e. el que tiene más "peso") se conoce como la *componente principal* de O . Obtener esta componente principal nos permite aproximar a O como una matriz de rango uno, y constituye una proyección, similar al proceso de minimización que se hace en el método de proyecciones generalizadas común [Kane, 2008]. Así, lo que se debe hacer para aplicar la restricción de forma matemática es obtener la componente principal de la nueva matriz señal obtenida en cada iteración, y los vectores cuyo producto externo corresponde esta componente serán los nuevos vectores \mathbf{E} y \mathbf{G} .

Es posible obtener esta componente principal descomponiendo O directamente en U , V y W , es decir, calcular directamente la descomposición en valores singulares. Sin embargo para el algoritmo PCGP solo son de interés los vectores que componen la componente principal, y computar toda la SVD, y por tanto las otras $N-1$ componentes, sería ineficiente. Afortunadamente existe una manera de obtener la componente principal sin tener que hacer toda la SVD, y se conoce como el *método de potencias*. Para usarlo pensemos no en los eigenvectores de O , sino en los eigenvectores de OO^T , llamemosles \mathbf{E}_i , y de $O^T O$, llamemosles \mathbf{G}_i . Éstos vectores cumplen:

$$\begin{aligned} OO^T \mathbf{E}_i &= \lambda_i \mathbf{E}_i, \\ O^T O \mathbf{G}_i &= \lambda_i \mathbf{G}_i, \end{aligned} \quad (4.6)$$

y se puede construir O como:

$$O = \sum_{i=1}^N \sqrt{\lambda_i} \mathbf{E}_i \mathbf{G}_i^T. \quad (4.7)$$

Es decir, O es la suma de los productos externos de los vectores \mathbf{E}_i y \mathbf{G}_i pesados con un

factor $\sqrt{\lambda_i}$, por lo que los vectores son las columnas de U y los renglones de V respectivamente. Notemos a continuación que si multiplicamos un vector arbitrario \mathbf{x} por OO^T , podemos pensar a ese vector arbitrario como una combinación lineal de los vectores \mathbf{E}_i , y entonces obtendremos:

$$OO^T \mathbf{x} = \sum_{i=1}^N \kappa_i \lambda_i \mathbf{E}_i . \quad (4.8)$$

Si multiplicamos continuamente el resultado por la matriz OO^T , digamos p veces, obtendremos:

$$(OO^T)^p \mathbf{x} = \sum_{i=1}^N \kappa_i (\lambda_i)^p \mathbf{E}_i . \quad (4.9)$$

Podemos notar que mientras p se vuelve más grande, el término con la λ más grande domina la suma completamente. Así, volviendo p grande se puede aproximar el vector \mathbf{E} que corresponde a la componente principal, y se puede hacer el proceso análogo para obtener el vector \mathbf{G} correspondiente. Así usando el método de potencias, se pueden obtener las nuevas propuestas del pulso señal y de la compuerta.

En la práctica, en vez de comenzar con un pulso \mathbf{x} arbitrario para aplicar el método de potencias y usar una potencia p grande, lo que se hace es multiplicar solo una vez, i.e. usar $p=1$, y usar como vector inicial la propuesta de pulso con la que se inició la iteración. Esto significa que en las primeras iteraciones no se estará obteniendo la componente principal de manera perfecta, pero entre más iteraciones se hagan, los vectores propuesta se parecerán más y más a los pulsos reales, y la matriz señal se parecerá más y más a la matriz señal real", que es de rango uno, y por lo tanto no sólo los vectores a los que se le aplica el método ya son muy similares a los de la componente principal, la matriz usada tiene una componente principal que domina sobre las otras. Es por esta razón que se encuentra en la práctica que usar $p=1$ es suficiente para que el algoritmo converja rápidamente [Kane, 2008] [Kane, 1998].

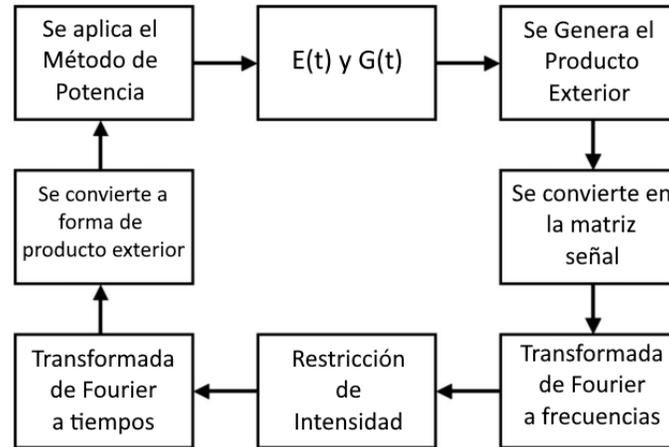


Figura 4.3: Diagrama de la estructura de una iteración del algoritmo FROG por PCGP, comienza y termina en el cuadro superior con la propuesta de una $E(t)$ y una $G(t)$

En resumen, en el algoritmo PCGP, después de aplicar la restricción de intensidad y transformar la matriz señal al espacio temporal, simplemente se invierte el proceso de rotación de renglones y manipulación de columnas que se hace al inicio de la iteración para poner la matriz en su forma de producto exterior, y posteriormente se multiplica el vector que se usó como propuesta inicial en esa iteración por el producto de la matriz con su transpuesta. El resultado de ese producto será el nuevo vector que se usará como propuesta en la siguiente iteración. Este procedimiento claramente es muy sencillo, especialmente comparado con una minimización de una distancia en un espacio de 256×256 , pues en realidad es solo una multiplicación de matrices, lo que hace que el algoritmo PCGP sea uno de los algoritmos FROG más veloces, suficientemente veloz incluso para hacer un sistema capaz de analizar trazas en tiempo real [Garduño-Mejía et al., 2003] [Kane, 1998]. En la figura 4.3 se muestra un diagrama de la estructura del algoritmo PCGP.

En el caso de que la traza sea una traza FROG SHG, donde la señal y la compuerta no sólo no son independientes, sino son iguales, hay dos opciones de como usar el algoritmo PCGP. La primera es dejar el algoritmo como se describió, con E y G independientes, y en principio el algoritmo debe converger en una solución en la cual sean iguales. La segunda opción es usar la información de que E y G no son independientes para modificar el programa, y forzarlo a que lo considere así. La manera de hacer esto es construir la matriz

señal al inicio de cada iteración no sólo como el producto exterior de \mathbf{E} y \mathbf{G} sino como la suma del producto exterior de \mathbf{E} con \mathbf{G} y el producto exterior de \mathbf{G} con \mathbf{E} . Esto fuerza a que los dos vectores que forman la componente principal de la matriz sean iguales [Kane, 2008].

Capítulo 5

Software en python

Para este trabajo se desarrolló un script en python que ejecuta el análisis de las trazas mediante el algoritmo FROG PCGP que se describe en el capítulo 4. El análisis requiere una cantidad considerable de poder de cómputo, una computadora doméstica común y corriente puede hacerlo sin mucho problema, pero un microprocesador del tipo que se podría incluir en el dispositivo mismo no es capaz de hacerlo en un tiempo razonablemente corto. Es por esto que el dispositivo solo obtiene las trazas, y el análisis se deja a una computadora externa al dispositivo. La idea del script es proveer junto con el dispositivo una manera conveniente de hacer el análisis para el usuario.

El script de análisis admite como entrada una imagen cuadrada en escala de grises de 8 bits. Se puede modificar en el script el tamaño de la imagen de entrada, para las pruebas se usaron imágenes de 128x128 y de 256x256, y da como resultado otra imagen de formato igual que representa la traza del pulso recuperado y una gráfica de los perfiles de intensidad y fase en el espacio de tiempo.

Para hacer pruebas con los scripts antes de que el dispositivo diera resultados reales se creó también un script sencillo para generar trazas simuladas a partir de un pulso dado. Éste script simplemente produce una traza a partir de la expresión matemática de la amplitud compleja de un pulso teórico. Genera un arreglo de puntos equidistantes en el

tiempo que luego evalúa en la expresión de la amplitud compleja y, usando la expresión de la ecuación 3.4, genera una matriz donde cada elemento, correspondiente a un par (t, τ) , representa la intensidad de un pixel de la traza simulada.

En las figuras 5.1 a 5.8 se muestran las trazas simuladas (a la izquierda) junto con las trazas correspondientes al pulso recuperado con el algoritmo. Se muestran también gráficas con los perfiles de intensidad y fase del pulso recuperado. Como se mencionó en los capítulos 3 y 4, la comparación de las trazas es una manera de evaluar la efectividad de la recuperación del pulso, pues una traza está asociada a un pulso de manera esencialmente única.

Durante las pruebas se usaron diferentes criterios para determinar si se ha llegado a la convergencia y romper el ciclo. La manera más obvia de hacerlo es calcular el error FROG como se definió en la ecuación 4.4 cada iteración, y considerar que se ha llegado a la convergencia cuando esté por debajo de un cierto número, que se tendría que definir mediante prueba y error. El cálculo se debe hacer después de crear la señal FROG en el espacio de retardo contra frecuencia, pero antes de comenzar a aplicar la restricción matemática. Éste método es muy ineficiente, pues calcular el error es un proceso intensivo en tiempo si se hace cada iteración. Tal como está definido el error en la ecuación 4.4 para calcularlo se requiere un paso de minimización para encontrar el factor μ . Es posible evitar la minimización normalizando ambas trazas al mismo máximo y usando μ igual a 1. El valor que se obtendrá será más similar al error FROG entre más parecidas sean las trazas, pero siempre será estrictamente mayor. Incluso así la adición de este cálculo en todas las iteraciones hace que el tiempo de cómputo para un número dado de iteraciones se multiplique varias veces. Otra opción es evitar hacer el cálculo hasta que hayan pasado cierto número de iteraciones, en las pruebas que se realizaron el algoritmo siempre tardó más de 30 o 40 iteraciones en converger satisfactoriamente, así que se puede realizar el cálculo del error solo después de las 40 iteraciones. Puede ser útil también hacer el cálculo solo cada cierto número de iteraciones, por ejemplo cada 5 iteraciones, y así se haría sólo en la iteración 45, 50, 55, 60, etc.

Otra manera de decidir si ya se ha llegado a la convergencia es calcular la diferencia en

el error de una iteración a otra, así se puede determinar cuando el algoritmo ha llegado a su límite, y no se necesita determinar un valor arbitrario del error en el cual se considera convergido. La versión actual del script usa este criterio, empieza calculando el error en la iteración 49, lo calcula también en la 50, los compara y repite el proceso cada 10 iteraciones. También se puede establecer un número determinado de iteraciones en las cuales terminar el proceso. En cualquier caso el programa reporta el error FROG, y la diferencia entre él y el de la iteración anterior como porcentaje.

Al igual que las trazas en la sección 3.1, las trazas que se presentan a continuación no tienen unidades. De la misma manera, los ejes en las gráficas del pulso recuperado no corresponden a unidades reales. El eje vertical representa intensidad normalizada a uno en el punto mas alto de intensidad del pulso ^[1] en las gráficas superiores, y fase en las inferiores, normalizada a π radianes. En ambos casos el eje horizontal representa el tiempo, en este caso en unidades arbitrarias que corresponden al incremento que representa un pixel en el eje horizontal de la traza. Una vez que se introduzca la calibración correspondiente al programa, este eje representará el tiempo en las unidades que se usen en la calibración.

Cabe mencionar que, como se mencionó en la sección 4.1 la traza no distingue factores constantes de fase ni desplazamientos en el tiempo. Por esta razón el pulso recuperado por el algoritmo normalmente tendrá la fase desplazada respecto al pulso original, pero ésto no constituye una distinción física real en un pulso, es equivalente a mover el cero en el eje del tiempo.

^[1]Cabe mencionar que este sería el caso aún con una traza real, pues no es práctico obtener la intensidad absoluta del pulso a partir de la luminancia en la imagen de la traza, lo importante en esta medición es el comportamiento de la intensidad respecto al tiempo, no su valor absoluto.

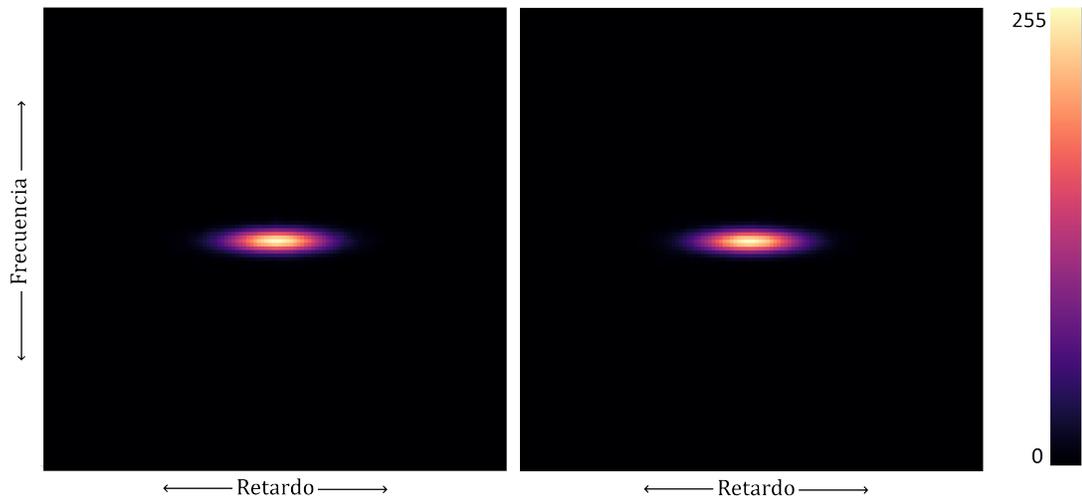


Figura 5.1: Trazas para un pulso con perfil de intensidad gaussiano y fase constante. Izquierda: Trazas simulada. Derecha: Trazas recuperada por el script de análisis.

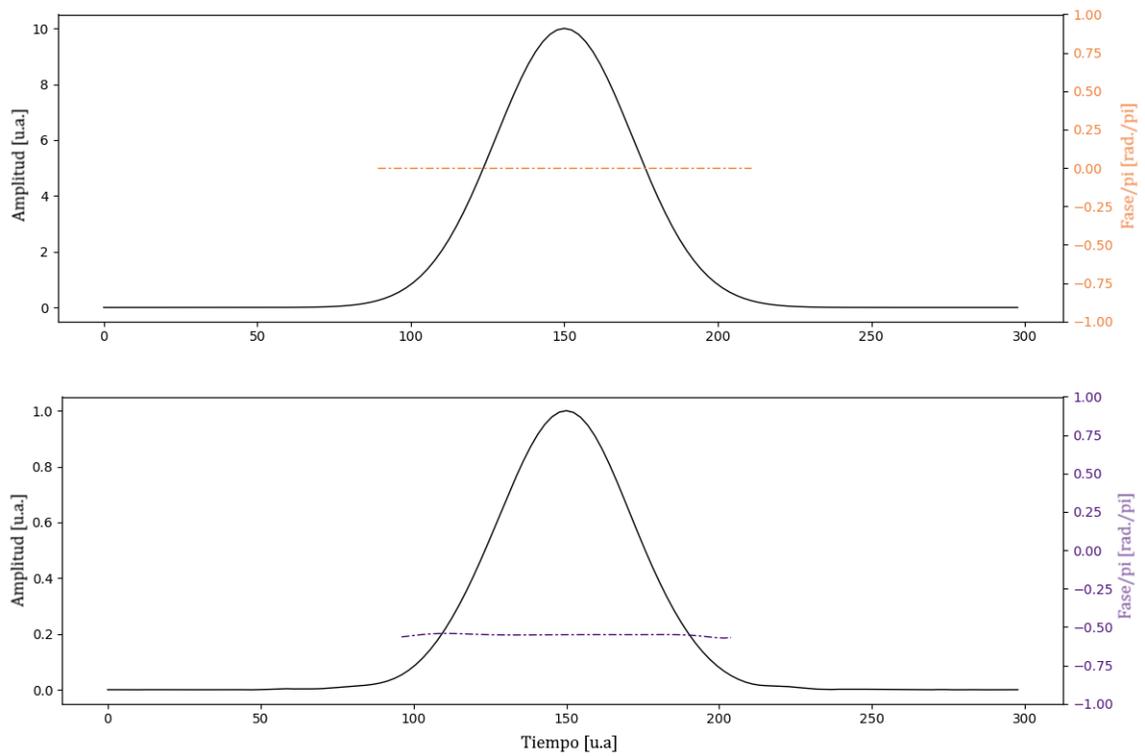


Figura 5.2: Amplitud y fase del pulso simulado y recuperado, un pulso gaussiano con fase constante. Arriba: Pulso simulado. Abajo: Pulso recuperado por el software.

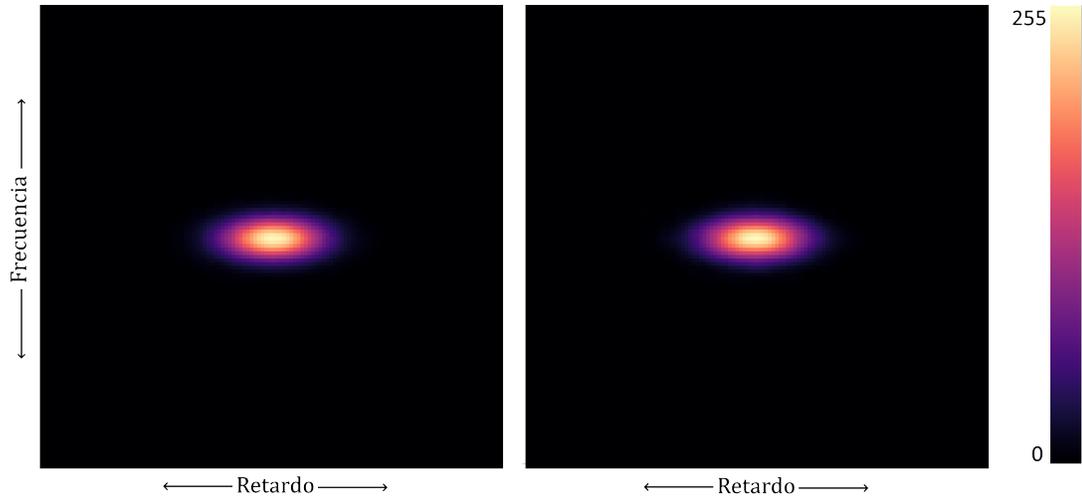


Figura 5.3: Trazas para un pulso con perfil de intensidad gaussiano y fase cuadrática. Izquierda: Trazas simulada. Derecha: Trazas recuperada por el script de análisis.

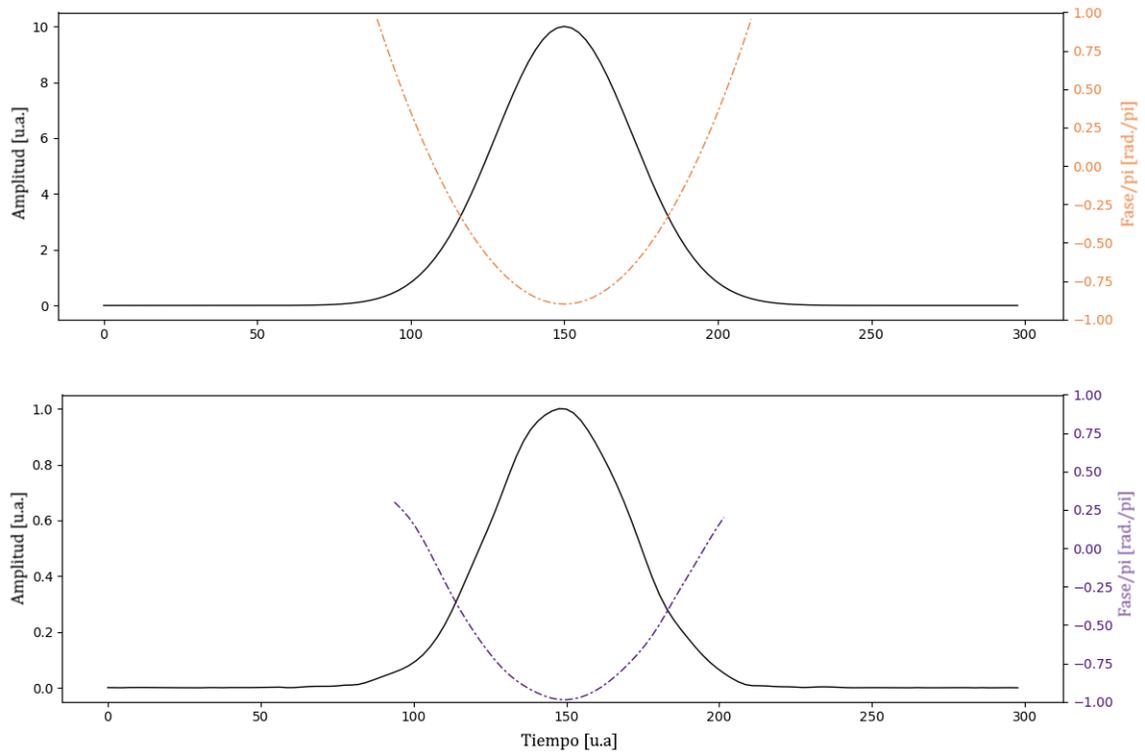


Figura 5.4: Amplitud y fase del pulso simulado y recuperado, un pulso gaussiano con fase cuadrática (i.e. chirp). Arriba: Pulso simulado. Abajo: Pulso recuperado por el software.

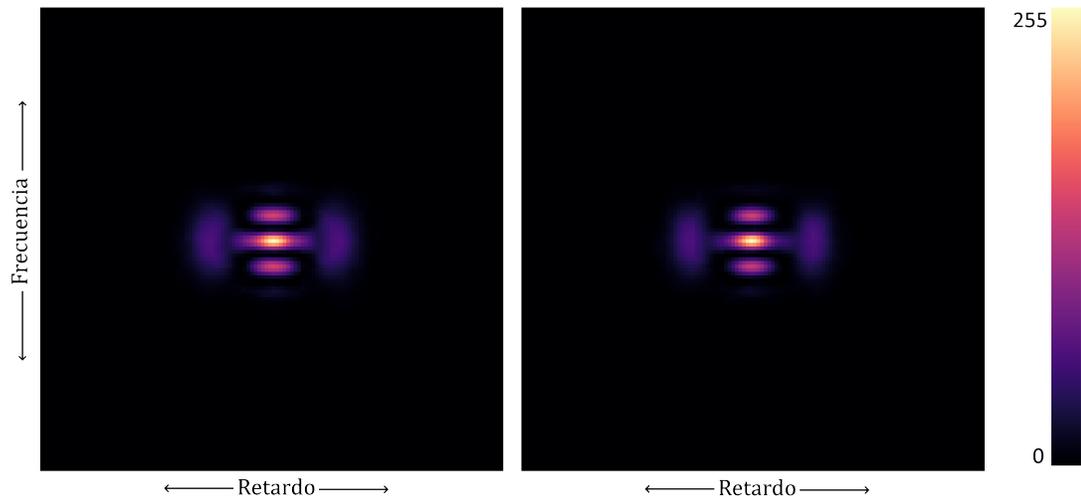


Figura 5.5: Trazas para un pulso con perfil de intensidad bimodal, hecha de dos perfiles gaussianos empalmados, cada uno con fase cuadrática. Izquierda: Trazas simulada. Derecha: Trazas recuperada por el script de análisis.

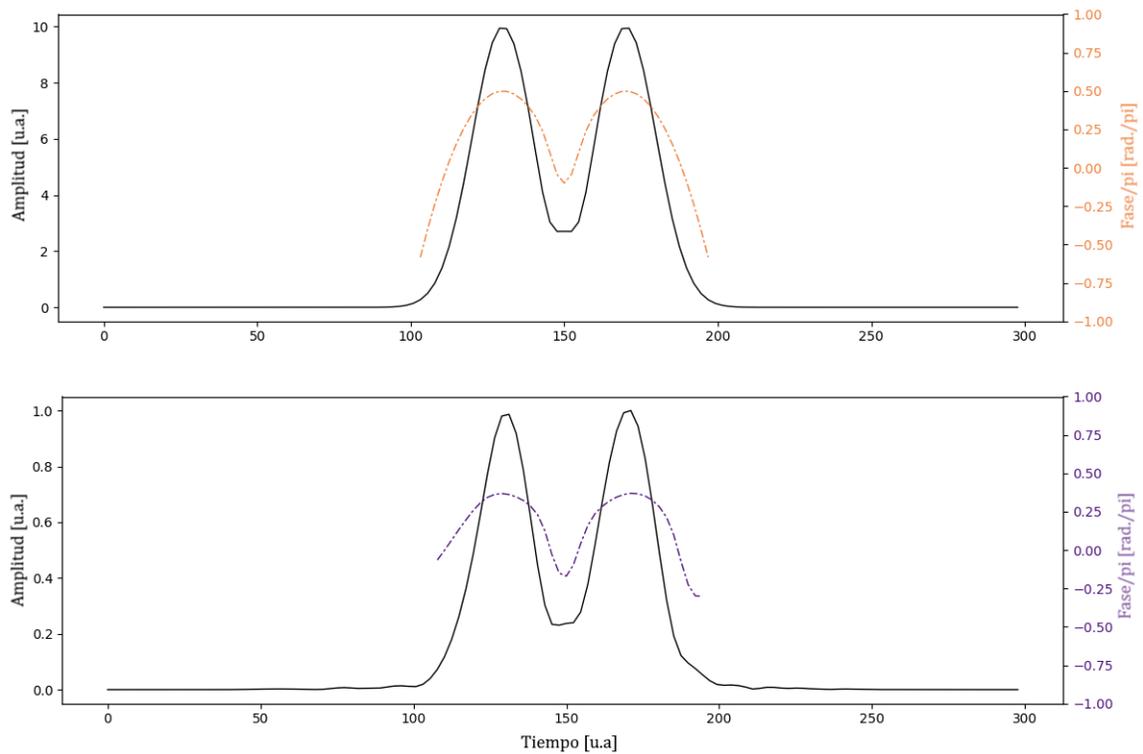


Figura 5.6: Amplitud y fase del pulso simulado y recuperado, un pulso con dos picos gaussianos, cada uno con fase cuadrática. Arriba: Pulso simulado. Abajo: Pulso recuperado por el software.

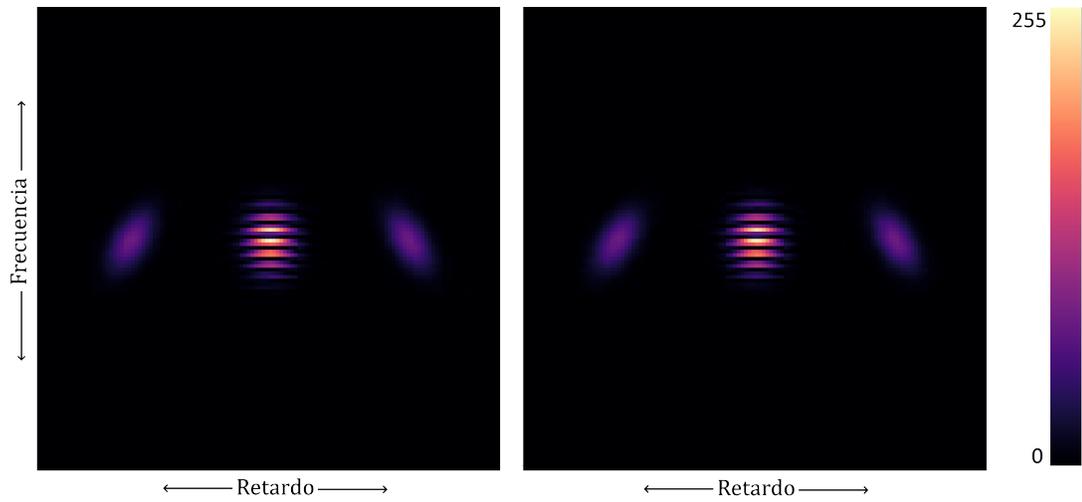


Figura 5.7: Trazas para un pulso doble, un tren de dos pulsos gaussianos con fases cuadráticas de distinto signo. Izquierda: Trazas simulada. Derecha: Trazas recuperada por el script de análisis.

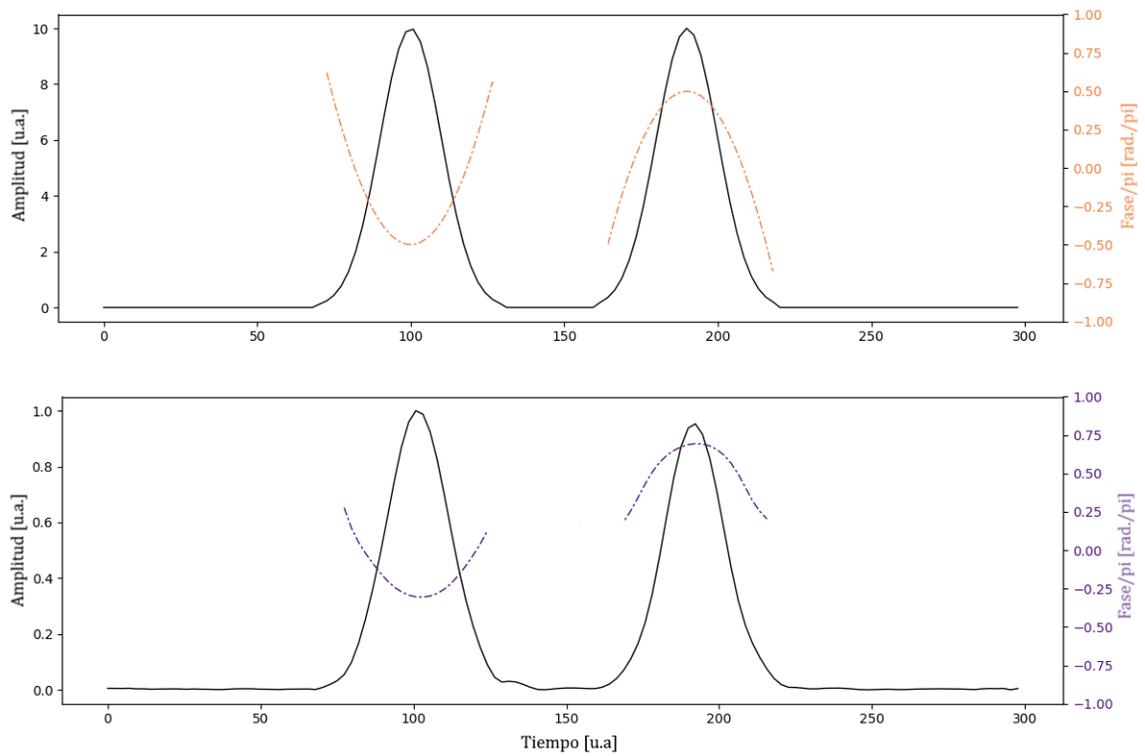


Figura 5.8: Amplitud y fase del pulso simulado y recuperado, tren de dos pulsos gaussianos con fases cuadráticas de signos opuestos. Arriba: Pulso simulado. Abajo: Pulso recuperado por el software.

La primera versión del script fue basada en la versión original del algoritmo frog, de la cual se habló en la primera sección del capítulo 4. Posteriormente se hizo una segunda versión, modificada para usar el algoritmo PCGP, tal como se describe en la segunda sección del capítulo 4. La primera versión, además de tardar más iteraciones en llegar a un resultado satisfactorio que la segunda, tenía una tendencia a estancarse y no converger en ciertos casos, por ejemplo en los pulsos con dos picos como los casos que se muestran en las figuras 5.5, 5.6, 5.7 y 5.8. La versión usando el algoritmo PCGP ha logrado converger para todos los casos probados.

El script está comentado y documentado con el propósito de ser una herramienta accesible a cualquier alumno con conocimiento básico de programación en Python. Se diseñó con fácil modificación y uso en mente, con la idea de que sea una herramienta flexible y abierta a modificaciones y mejoras en el futuro. Uno de sus límites en su versión actual es su eficiencia. Como ya se mencionó el algoritmo usado es altamente eficiente, sin embargo puede ser que la implementación no sea óptima. La versión actual corriendo en una laptop doméstica^[2] puede completar aproximadamente 3 iteraciones por segundo. Otra modificación que podría hacerse en el futuro para volverla una herramienta más accesible sería hacer una interfaz gráfica.

^[2]del año 2014, con 8GB de RAM y procesador intel core i5 de 1.7GHz

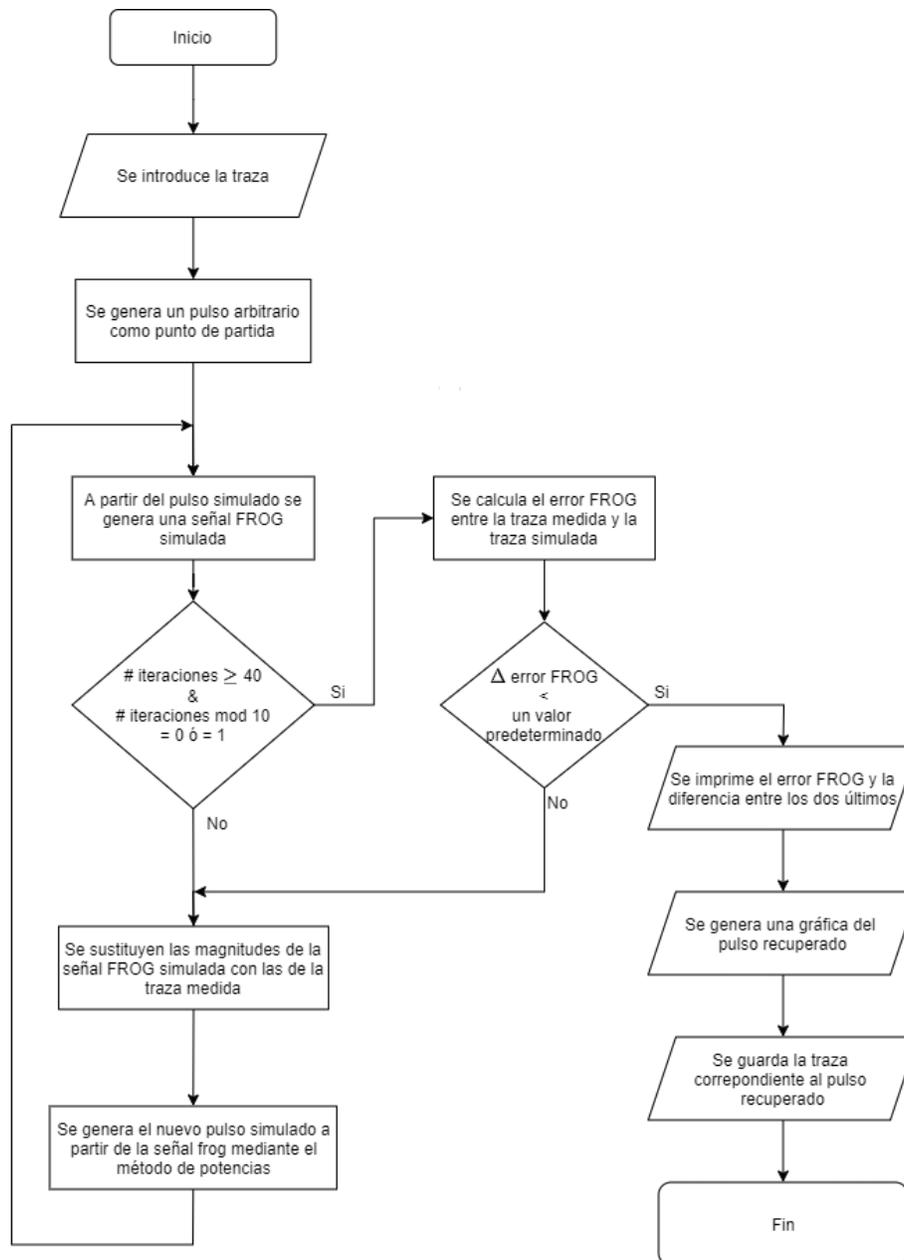


Figura 5.9: Diagrama de flujo que representa el funcionamiento del script de análisis.

Capítulo 6

Diseño del dispositivo

Como se mencionó anteriormente en la introducción general, el dispositivo diseñado en este trabajo se basó en un dispositivo que ya existía en el laboratorio de pulsos ultracortos del ICAT, y que fue diseñado y construido por un alumno de maestría. El propósito de este trabajo fue rediseñar y modificar el aparato para mejorar su funcionamiento. De esta versión anterior del dispositivo se heredó la base principal en la cual se monta el arreglo óptico, y se usaron los mismos elementos ópticos (o elementos equivalentes) excepto en la sección del espectrómetro, en la cual hubo fuertes modificaciones (que se discutirán a detalle más adelante).

El problema principal del dispositivo existente era su velocidad de adquisición. En vez de usar un espejo rotatorio para barrer la cámara y generar el eje correspondiente al retardo en la traza, como se muestra en el diagrama 3.3, simplemente tomaba una imagen del espectro para cada posición de la línea de retardo. Posteriormente se usaban todas estas imágenes para generar la traza digitalmente, extrayendo una sección de cada imagen y juntándolas en orden como se muestra esquemáticamente en la figura 6.1[Kapellmann, 2012]. Esto simplifica el diseño óptico y facilita el control del aparato, pues no es necesario sincronizar dos partes móviles. Sin embargo el hecho de que se debe tomar una imagen por cada columna de píxeles en la traza final, y el hecho de que se debe hacer un procesamiento no trivial de todas estas imágenes para obtener la traza hacen que el proceso sea

súmamente tardado e ineficiente. Éste era el principal problema a resolver. Por esta razón se tuvo que reemplazar todo el sistema de control y el software, que estaban intrínsecamente diseñados para trabajar de la manera anteriormente descrita.

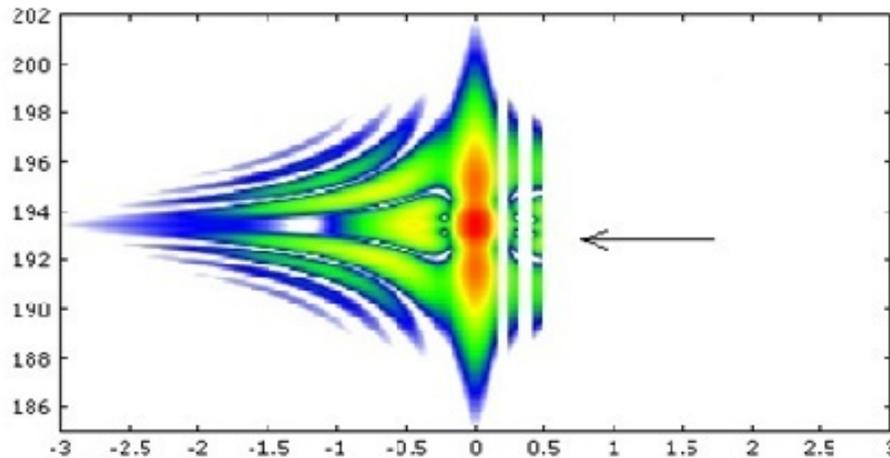


Figura 6.1: Esquema de cómo el software del dispositivo precursor junta los espectros en cada retardo para formar la traza.

En este capítulo se describirán tanto los elementos que se heredaron directamente del dispositivo precursor como los elementos que se agregaron, señalando en cada caso si fueron modificaciones y el razonamiento de la modificación. Se describirá el aparato en tres secciones. Como ya se mencionó en el capítulo 3, el arreglo óptico consiste esencialmente de dos partes, un arreglo autocorrelador y un espectrómetro. La tercera sección tratará sobre el diseño del sistema de control.

6.1. El autocorrelador

El autocorrelador se mantuvo esencialmente sin cambios respecto al dispositivo precursor, sin embargo será importante que se describan ciertos aspectos de su diseño. Su diseño es como se muestra en la figura 3.2, se compone de tres espejos, un divisor de haz y dos lentes que enfocan los haces en un cristal no lineal y los coliman posteriormente. Se dejará la discusión de la línea de retardo a su propia subsección, la subsección 6.1.1. En la

figura 6.2 se muestra el autocorrelador armado, excepto por el cristal, durante el proceso de alineación.

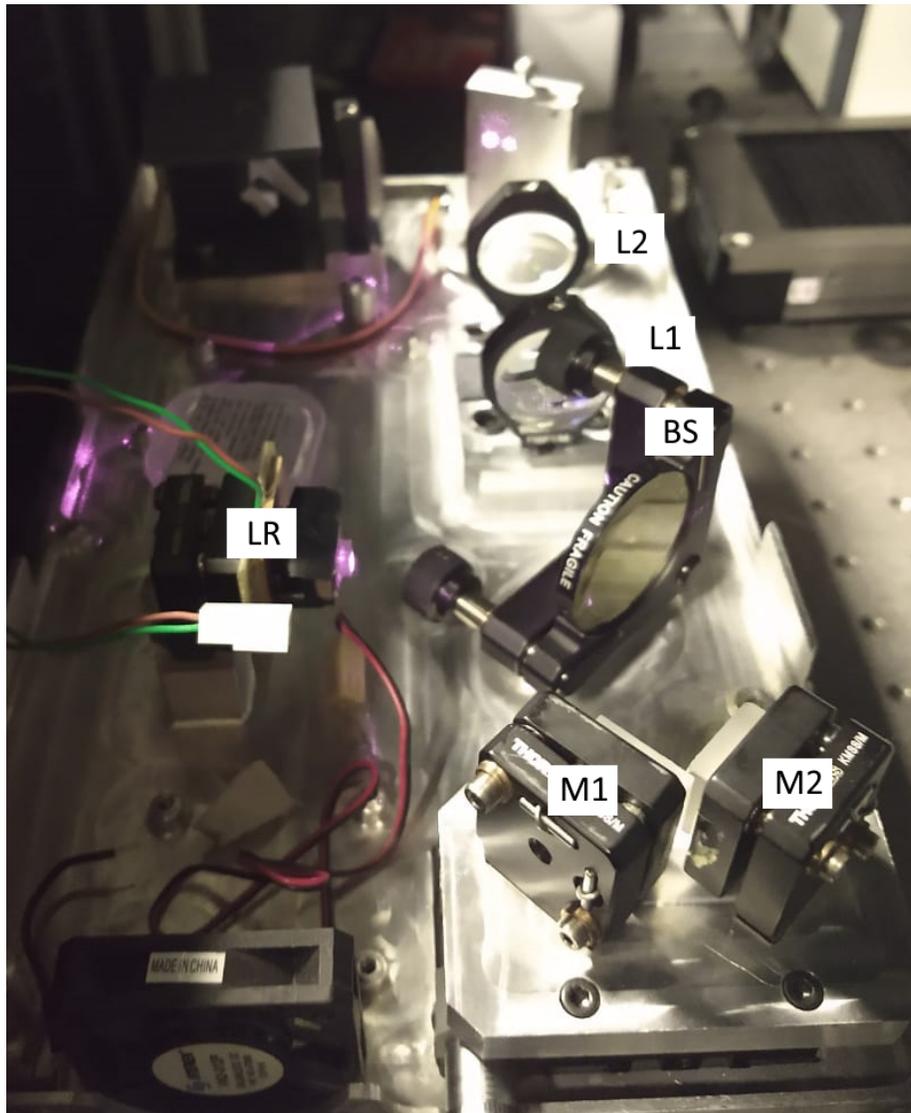


Figura 6.2: Foto del arreglo durante el proceso de alineación, antes de poner el espectrómetro y el sistema de control. Se etiquetan los elementos de la misma manera que en las figuras 3.2 y 3.3

El dispositivo FROG fue probado durante este trabajo usando el láser Mira Optima 900-F, un láser pulsado de Ti:Zaf con longitud de onda de alrededor de 800nm, que se encuentra en el laboratorio de pulsos ultracortos del ICAT. El dispositivo FROG es portátil, debe poder moverse y usarse con otros láseres pulsados, sin embargo hay consideraciones que se deben de hacer sobre la longitud de onda y duración de pulso de un láser para

saber si el dispositivo puede ser usado para medirlo. En lo que respecta a la longitud de onda, el recubrimiento de las lentes y el divisor de haz están optimizados para el infrarrojo cercano, si se usan con un láser de otra longitud de onda podría haber reflejos no deseados que causen pérdidas. Sin embargo la mayor restricción sobre la longitud de onda que se quiere medir es el cristal no lineal usado en el autocorrelador, pues como se mencionó en el capítulo 2, estos cristales están diseñados para usarse a longitudes de onda específicas. Si se deseara medir un láser con una longitud de onda lejana a los 800nm, se necesitaría usar un cristal diferente. Sin embargo el resto del arreglo funcionaría completamente como debe.

Una observación interesante sobre el láser usado para las pruebas y la alineación es que, como ya se mencionó, tiene una longitud central de 800nm, su espectro contiene un poco de rojo visible pero en general no es visible al ojo humano, y sin embargo en la cámara se ve y aparenta ser de un color magenta, como se ve en la figura 6.2^[1]. Esto fue útil para el proceso de alineación, ya que a veces resultó más conveniente usar el celular como visor infrarrojo que los visores infrarrojos que hay en el laboratorio.

Los espejos M1 y M2 están en un arreglo a 90 grados para trasladar el haz y que no sea colineal con el haz que va a la línea de retardo. Se encuentran sobre una platina de traslación, para poder modificar la longitud de camino óptico. Dado que la línea de retardo tiene un rango pequeño de movimiento se necesita poder modificarla en este brazo del arreglo para encontrar la posición donde la longitud de camino óptico sea igual en los dos brazos.

Para dividir el haz entrante se usó un divisor de haz de película delgada. Este tipo de divisor está diseñado para minimizar los efectos de dispersión y así no afectar la forma de un pulso ultracorto. En otras partes del diseño óptico se usaron lentes comunes, que como cualquier medio tienen dispersión, y pueden afectar la forma del pulso. El efecto dependerá de la duración del pulso, entre más corto sea el pulso mayor será, en proporción, el

^[1]Esto se debe a que los fotosensores que componen una cámara digital, que están diseñados para sensar luz roja, verde y azul, también son sensibles a la luz infrarroja cercana al visible. Usualmente las cámaras tienen un filtro para tapar la luz infrarroja, pero en el caso de cámaras baratas o muy básicas (como lo son las de los celulares) o en casos donde la luz infrarroja es de muy alta intensidad, el filtro no logra bloquear toda. Evidentemente en este caso los sensores rojos y azules son más sensibles al infrarrojo que los verdes, y por tanto el láser aparece magenta o rosa

efecto de la dispersión sobre él. Como ejemplo, el vidrio BK7 presenta una dispersión de velocidad de grupo de $44.7 fs^2/mm$ [Schott-AG, 2017]. Si un pulso gaussiano con duración de $100 fs$, inicialmente sin chirp, atraviesa una longitud de $10 mm$ de BK7 su duración se elongará a $104 fs$. Por otro lado un pulso de $30 fs$ se elongará a $42 fs$ al atravesar el mismo material. El efecto que tendrá la dispersión depende también de el chirp que tenga inicialmente el pulso [Diels and Rudolph, 2011]. Si se quisiera eliminar el problema de la dispersión, y así permitir que el aparato pueda medir pulsos más cortos sin deformarlos, se podría modificar el módulo en el que se montan las lentes y el cristal para usar espejos curvos en vez de lentes para enfocar y colimar el haz. De esta manera el pulso no tendría que atravesar un material dispersivo (salvo por el aire).

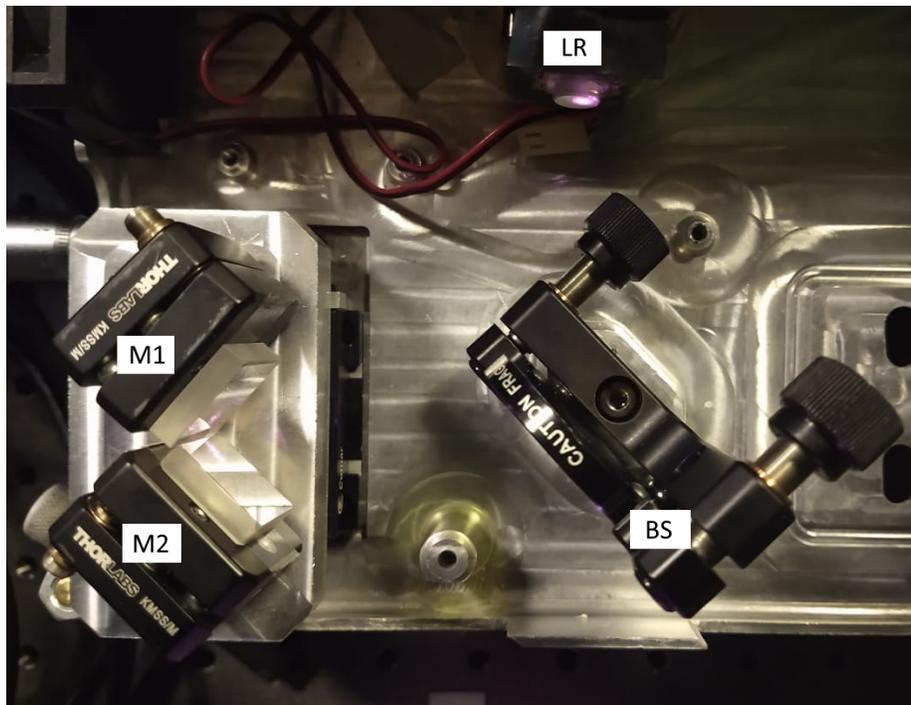


Figura 6.3: Foto de la sección inicial del autocorrelador, donde se observan los espejos M1 y M2, el divisor de haz BS y la línea de retardo LR.

Las lentes están montadas sobre un módulo que permite modificar su posición para asegurar que estén alineadas con el camino de los haces. Están acomodadas a una distancia igual al doble de su distancia focal (y también se puede modificar y afinar la distancia entre ellas), de tal manera que enfocan y juntan los haces cuando vienen paralelos, y los colima y los hace paralelos al salir. Este módulo tiene espacio para fijar y mover el cristal no lineal.

La pieza del módulo que sostiene la montura del cristal se rediseñó para compensar que la parte de la montura donde se coloca el cristal no está en el mismo eje que el tornillo con el cual se agarra la montura. Se diseñó el modelo 3D de la pieza en solidworks y posteriormente se usó una impresora 3D para imprimir la pieza. Es una pieza muy simple, rectangular y con dos esquinas redondeadas, dos hoyos en la parte superior para atornillar la montura y uno para fijarla al módulo. Se muestra el modelo de la pieza en la figura 6.4.

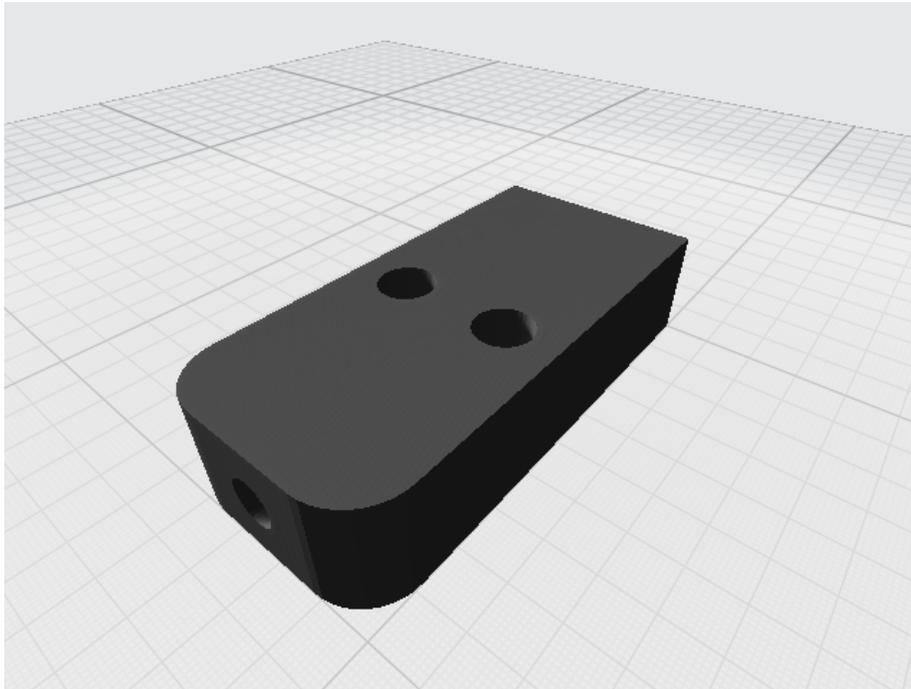


Figura 6.4: Modelo 3D de la pieza que sostiene la montura del cristal no lineal.

Se usó en este trabajo un cristal BBO hecho para generar SHG por phase matching tipo 1 con un grosor de 200 micras. El cristal se monta en un soporte que puede rotarlo alrededor del eje que definen las lentes para encontrar la orientación correcta.

6.1.1. La línea de retardo

En las figuras 6.2 y 6.3 se puede ver el elemento que compone la línea de retardo. Es un actuador lineal obtenido de una webcam, era parte del sistema de enfoque y es sencillamente un solenoide al cual se le pegó un pequeño espejo. Al pasar corriente por una bobina que rodea al elemento, una parte de éste se mueve en función de la cantidad de corriente que se le introduzca. El actuador es sumamente compacto, ligero, y se puede controlar electrónicamente de manera relativamente sencilla sin necesidad de impartirle mucha corriente. El uso de este elemento fue heredado del dispositivo anterior. Fue usado anteriormente en un trabajo publicado en *Review of Scientific Instruments*, por miembros del laboratorio de pulsos ultracortos del ICAT, como línea de retardo en un autocorrelador por absorción de 2 fotones [Castro-Marín et al., 2015]. Una de las ventajas que tiene es su bajo costo, pues es parte de un aparato comercial muy común y por tanto se puede obtener incluso de webcams desechadas, esencialmente de la basura, sin ningún costo. También es un elemento sencillo de controlar, simplemente su posición depende de la corriente que se pase a través del elemento (más adelante se verá que esa dependencia es aproximadamente lineal, al menos en parte del rango de movimiento), y basta solo un poco de electrónica para usarlo para este propósito. En la sección 6.3 se describirá a detalle cómo se controla este elemento.

Se realizaron dos pruebas independientes para caracterizar el movimiento de la línea de retardo. En el trabajo del dispositivo original se hizo una caracterización muy exhaustiva del mismo componente [Kapellmann, 2012], sin embargo dado que se modificó la electrónica que se usa para manejarlo y dado que ha pasado algún tiempo desde entonces, se decidió caracterizar el dispositivo nuevamente. La primera de estas pruebas se efectuó usando interferometría. Al poner la línea de retardo en el espejo en uno de los brazos de un interferómetro de Michelson, y medir la intensidad a la salida mientras se mueve a velocidad constante, se puede usar la interferencia observada para determinar la velocidad a la que se mueve la línea de retardo. Se midió al mismo tiempo el voltaje introducido y así se puede determinar como corresponde el voltaje con la posición. Se muestran en las figuras 6.5 y 6.6 imágenes de la medición en el osciloscopio.

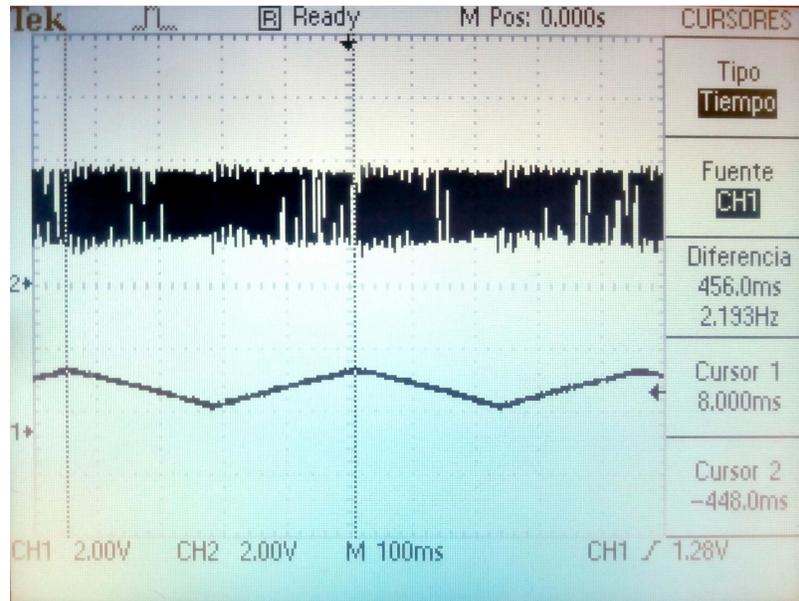


Figura 6.5: Arriba se muestra la señal del interferómetro. Abajo se muestra la señal introducida a la línea de retardo, una rampa que sube y baja a velocidad constante en cada segmento.

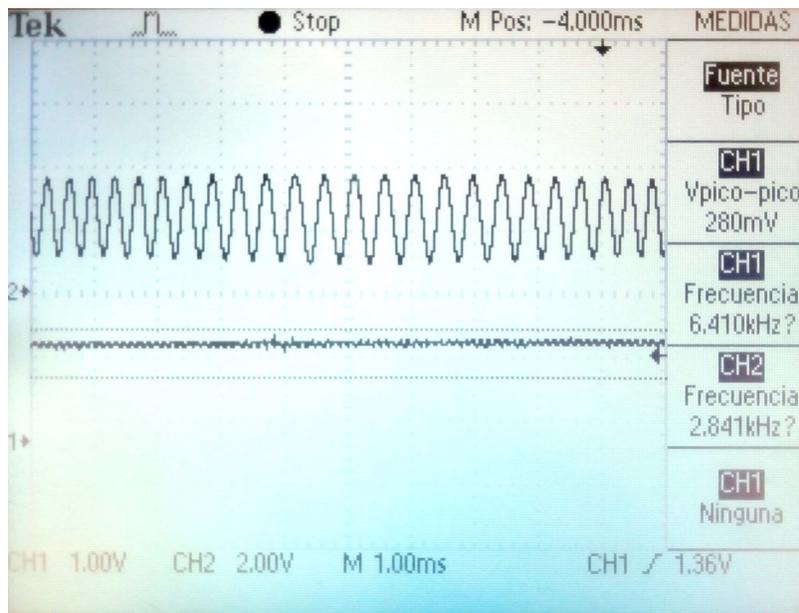


Figura 6.6: Arriba se muestra la señal del interferómetro en una escala más cercana a la figura anterior, donde se observan las franjas causadas por la interferencia. Abajo se muestra la señal introducida a la línea de retardo, a esta escala no se observa su pendiente.

Se descubrió al hacer esta prueba que, dado que el actuador no está diseñado para moverse rápidamente, si se trata de moverlo hacia adelante y hacia atrás rápidamente el espejo tiende a oscilar lateralmente, lo que desalinea el arreglo. En la figura 6.7 se ve una oscilación en la intensidad de la señal que indica que el espejo oscila, sale y entra en alineación con el resto del arreglo. Ésto impone un límite a la velocidad a la que se pueden obtener trazas. Al hacer la prueba se aseguró que la línea de retardo se moviera lo suficientemente lento para no generar estas oscilaciones en el espejo.

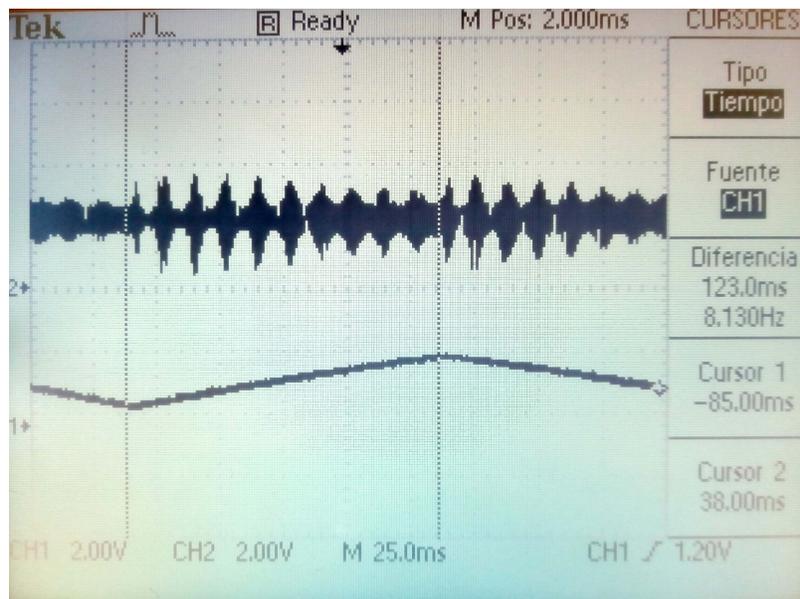


Figura 6.7: Arriba se muestra la señal del interferómetro, se observa la oscilación de su intensidad. Abajo se muestra la señal introducida a la línea de retardo.

Se midió la frecuencia de la señal del interferómetro a lo largo de cada subida y bajada del voltaje introducido a la línea. La línea de retardo se controla mediante una señal PWM (*pulse width modulation* o modulación por ancho de pulso), que se pasa por un filtro simple para asemejarla a una rampa de voltaje continua. Se describirá a detalle el control de la línea en la sección 6.3. La variable que se usa para controlar el ciclo de trabajo de este PWM es un número de 8 bits, es decir, un entero entre 0 y 255. Se usará este número para describir el voltaje efectivo introducido a la línea, dado que es la variable que se controla directamente. La rampa de voltaje que se usó en la prueba anteriormente descrita va de un valor de 60 a un valor de 180, correspondiente a ciclos de trabajo de 23 % y 70 %

respectivamente, en 0.228 segundos. Se observó que el movimiento era consistente dentro de este rango, aunque no fue hasta la segunda prueba, que se describirá más adelante, que se determinó realmente en qué rango el movimiento es lineal. Midiendo en varias rampas, la señal del interferómetro mostró una frecuencia en promedio de 2.6KHz, con una desviación estándar de 0.27KHz. Considerando que se usó un láser de He-Ne con una longitud de onda de 644nm, se puede calcular fácilmente que el espejo se mueve $1.57 \pm 0.17\mu m$ por cada cambio de un valor entero del ciclo de trabajo, considerando el rango que se usó anteriormente, entre 60 y 180, se tiene un rango de movimiento de la línea de retardo de aproximadamente 190 μm . Además de determinar el rango, esta prueba confirma que el movimiento de la línea de retardo es suave y se usó para determinar a qué velocidad se puede mover sin causar problemas de alineación.

Como se mencionó en el capítulo 3, el propósito de la línea de retardo es modificar el retardo relativo entre los dos haces, retardando uno de los dos pulsos que llegarán al cristal no lineal juntos. Debe asegurarse en la práctica que el rango de movimiento de la línea de retardo sea suficiente para obtener toda la información de la autocorrelación, lograr al menos que se pueda modificar el retardo relativo lo suficiente para que cada punto de uno de los pulsos haya interactuado con cada punto de el otro. Se puede considerar que para lograr obtener toda la información necesaria la línea de retardo se debe mover una distancia correspondiente en tiempo a tres veces el ancho del pulso a medir^[2]. Si se considera el rango de movimiento encontrado en la prueba mencionada anteriormente, de 190 μm , eso corresponde a un rango temporal de 633fs, y por lo tanto se podrían medir pulsos de hasta 210fs usando este rango.

Una vez que se completó la sección del autocorrelador del dispositivo, se realizó una segunda prueba para determinar mejor el rango en el que el movimiento de la línea de retardo es lineal. Como se mencionó anteriormente en la sección 6.1, el brazo fijo del autoco-

^[2]Si uno imagina la longitud "total" de un pulso, es fácil ver que partiendo desde el punto donde las dos copias del pulso se empiezan a traslapar, un cambio de una longitud los empalmará completamente, y un cambio de dos longitudes los llevará al punto donde terminan de traslaparse. Dado que al hablar de la longitud de un pulso normalmente usamos medidas como el ancho FWHM, que no cubre todos los puntos donde la intensidad del pulso es distinta de cero, se necesita un recorrido un poco mayor. Como regla general, entonces, se puede considerar que tres veces la longitud es una estimación razonable de la distancia con la cual podemos asegurar que se verá claramente la autocorrelación completa. Evidentemente esto depende del pulso particular, pero esta estimación resulta útil para establecer las limitaciones del arreglo.

rrrelador ^[3] tiene longitud ajustable mediante un tornillo micrométrico, para poder igualar las longitudes de ambos caminos ópticos en el centro del rango fijo de la línea de retardo. Se usó esta característica para hacer la segunda caracterización de la línea de retardo. El procedimiento fue el siguiente: primero se alineó el autocorrelador, de manera que los dos brazos tuvieran la misma longitud de camino óptico, y por lo tanto se estuviera en la condición de retardo cero, donde el autocorrelador da su máxima intensidad. Se aseguró que la intensidad fuera máxima midiendo la intensidad del haz con un fotodetector. Posteriormente se modificó el voltaje aplicado a la línea de retardo (modificando el ciclo de trabajo del PWM), y se ajustó la posición del otro brazo con el tornillo micrométrico hasta volver a llegar al máximo de intensidad. Esto asegura que ambos brazos se han movido la misma distancia, y el tornillo mide cuánto se ha movido. De esta manera se puede saber la posición de la línea de retardo para cada valor de voltaje aplicado.

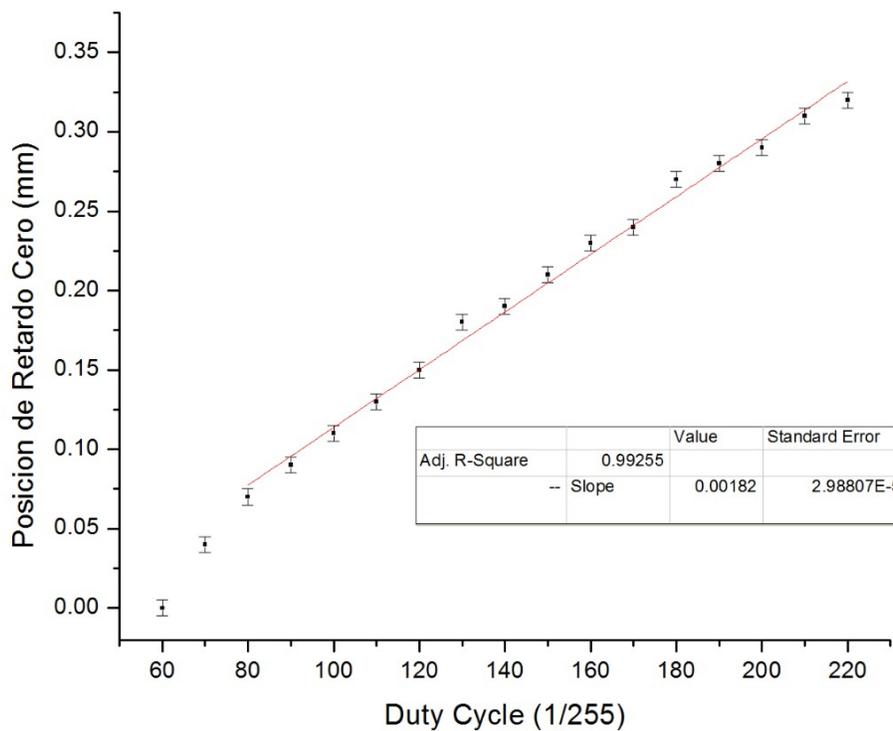


Figura 6.8: Gráfica que representa el movimiento de la línea de retardo respecto al valor del ciclo de trabajo del PWM que la controla. Se le ajusta una recta en el rango de 80 a 220

Se muestra en la figura 6.8 la gráfica resultante de esta prueba. Para valores muy ba-

^[3]es decir, el que no tiene la línea de retardo

jos de voltaje, el actuador descansa en una orientación ligeramente distinta a los valores más grandes ^[4], lo que desalinearía el arreglo si usáramos esos valores durante el funcionamiento del aparato. Es por esto que el eje horizontal de la gráfica comienza a contar en 60, que corresponde a un ciclo de trabajo de 23 %. Se observa que a valores bajos el comportamiento no es lineal, se curva hacia abajo, sin embargo de un valor de 80 en adelante el comportamiento se aproxima bien por una recta, con pendiente que corresponde a $1.82 \pm .03 \mu m$ por cada valor entero del ciclo de trabajo. Este resultado no coincide con el encontrado en la prueba con el interferómetro, cae ligeramente fuera de la incertidumbre. Dado que esta es una medición más directa del funcionamiento como tal en el aparato final, se usará lo obtenido en esta prueba. Si se usara todo el rango que se ajustó a la recta, significaría un rango de movimiento de 250 micras, que corresponde a un rango temporal de aproximadamente $833 fs$, lo suficiente para medir pulsos de hasta $280 fs$.

6.2. El espectrómetro

Después de la sección que contiene las lentes y el cristal se usa un diafragma para separar los haces fundamentales del haz de suma de frecuencias, y comienza la sección del espectrómetro. Se coloca una rejilla de difracción de 1500 líneas por milímetro, montada en una pieza circular de tal manera que pueda rotar sobre el eje vertical para ajustar la dirección en la que se dirige el haz.

La sección posterior a la rejilla fue modificada considerablemente. Como ya se mencionó al inicio de este capítulo, en este trabajo se añadió un elemento completamente nuevo, el espejo rotatorio. Dado que el espacio en la base estaba diseñado para el arreglo original, sin el espejo rotatorio, se diseñó una montura en solidworks para contenerlo y se imprimió una pieza tridimensional. Fue diseñada para acoplarse con la montura ya existente para la lente L3 y la cámara. Para introducir el espejo rotatorio sin requerir más espacio de la base, se diseñó la montura para desviar el haz en dirección vertical, con el módulo CCD apuntando hacia abajo. Se puede ajustar la posición de la montura para asegurar que la

^[4]probablemente porque en estos valores la parte móvil comienza a tener contacto con la coraza que cubre el resto del elemento

cámara esté en el plano focal de la lente. Se muestra el modelo 3d en la figura 6.9, y se muestra la pieza montada en el arreglo en la figura 6.10, con el espejo y la cámara.

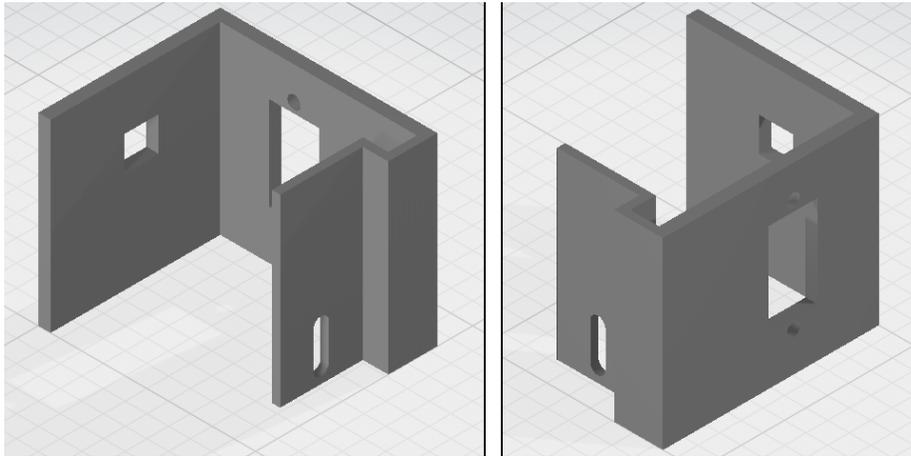


Figura 6.9: Modelo 3d de la pieza diseñada para sostener el espejo rotatorio y la cámara. Se muestran dos perspectivas.

La pieza se monta con un tornillo a través de una ranura, de manera que permite ajustar la posición. El agujero rectangular fue diseñado para montar uno de los servomotores que se mencionarán en la sección 6.2.1.

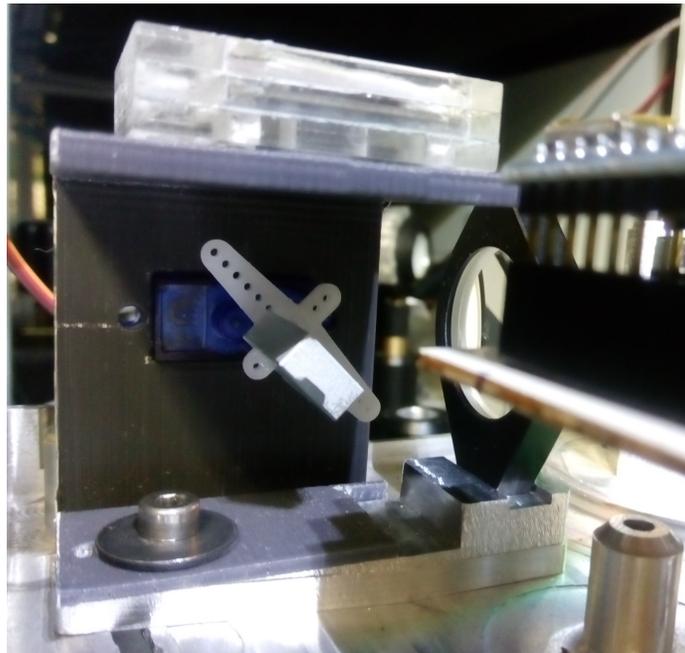


Figura 6.10: Pieza impresa en 3D y montada en el arreglo, junto con el espejo rotatorio, la lente L3 y la cámara.

Cuando el dispositivo está en uso se cubre toda la sección mostrada en la figura 6.10 con cartulina negra en los lados que tiene abiertos para evitar que entre luz externa a la cámara.

6.2.1. Espejo rotatorio y cámara

Para el espejo rotatorio se consideraron varias opciones y se hicieron pruebas con dos versiones del elemento. La idea es tener un espejo montado en algún tipo de motor que lo haga rotar sobre un eje contenido en el plano donde está la superficie reflectora del espejo de manera simultánea al movimiento de la línea de retardo. Se usaron durante el proceso del diseño del dispositivo dos motores diferentes, inicialmente para prototipar el sistema se usó un servomotor SG-90, pequeño y de bajo costo. Se encontraron ciertas fallas con el uso de este motor que se describirán a detalle en el capítulo 7. Posteriormente se probó un motor de pasos NEMA-17 a través de controladores capaces de usar fracciones del paso normal del motor, la documentación del controlador se refiere a estas fracciones como micropasos. Es necesario usar micropasos pues el módulo ccd tiene un ancho de aproximadamente 5mm y queda a una distancia de aproximadamente 22mm del eje del espejo, lo que significa que el rango completo de movimiento angular del motor durante el funcionamiento debe ser de aproximadamente 12° como cota máxima. Dado que cada paso del motor es de 1.8° , sin el uso de micropasos habría menos de 10 pasos para recorrer la cámara, y se necesita que el movimiento sea lo más suave posible, hecho de la mayor cantidad de pasos posible. En la sección 6.3 se hablará más a detalle de los controladores que se usaron para el motor a pasos y su comportamiento.

A ambos motores se les añadió un pequeño espejo de superficie metálica, pegado de tal manera que su superficie reflejante coincida con el eje de rotación del motor. El servomotor se montó directamente en la pieza mencionada en la sección anterior, y que se muestra en la figura 6.10. Por otro lado, dadas las dimensiones del motor a pasos, por el momento se necesita un soporte separado para sostenerlo y acomodar su eje.

En la parte superior de la montura impresa en 3D se coloca el módulo CCD. CCD

(*charge-coupled device*) se refiere a un tipo de circuito integrado que consiste de un arreglo de condensadores acoplados con el cual se puede sensar la luz que incide en un área. Se usan comúnmente como sensores para cámaras digitales. En este caso se trata de un módulo de cámara de raspberry al que se le removió la lente, de manera que el sensor queda expuesto. El módulo está contenido en una pequeña caja de acrílico (que se ve en la figura 6.10) con un hueco para que entre la luz al sensor. Cuando no está colocado o en uso se debe tapar para evitar que entre polvo al sensor, ya que no se puede limpiar fácilmente. Es importante que la cámara esté colocada en el plano donde enfoca la lente. Para este propósito se debe ajustar la posición de la montura de adelante hacia atrás. Es útil tomar imágenes del espectro que llega a la cámara sin mover la línea de retardo ni el espejo, y asegurar que se vea una línea delgada y clara como se muestra en la figura 6.11. También es importante asegurar que no entre mucha luz externa a la cámara. Para hacer estas pruebas se debe modificar el tiempo de exposición del sensor. Es importante considerar que durante el uso normal se usará un tiempo de exposición muy largo, lo que puede empeorar considerablemente el ruido causado por la luz externa. Se hablará a más detalle sobre el control de la cámara en la sección 6.3.

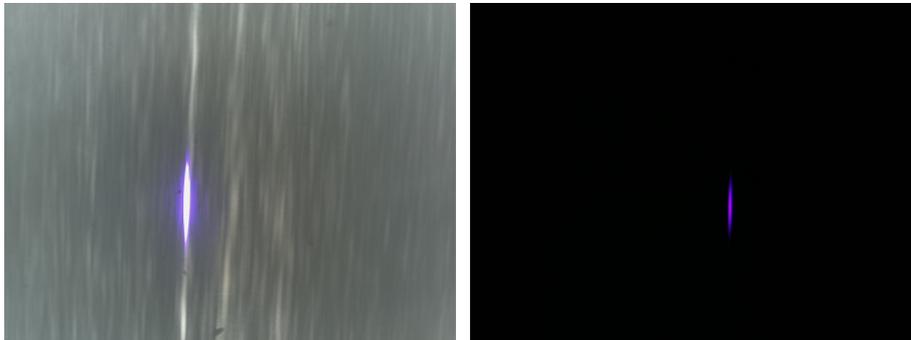


Figura 6.11: Comparación entre dos espectros tomados con la cámara. A la izquierda uno con mal enfoque y con luz de fondo. A la derecha uno enfocado correctamente y sin luz de fondo.

6.3. Sistema de control

Como se mencionó al inicio de este capítulo, se cambió completamente el sistema de control. El dispositivo es controlado principalmente mediante un Raspberry Pi. Esto permite tener una interfaz gráfica con el aparato y visualizar las trazas antes de transferirlas a una computadora. El dispositivo se puede conectar con una computadora mediante una conexión VNC^[5] para controlar, visualizar y mover archivos remotamente a través de un cable ethernet o por wifi^[6]. Esto también hace que el dispositivo sea más accesible y sencillo de usar y modificar, ya que se puede leer y modificar el programa de control en el mismo dispositivo. El Raspberry controla directamente el módulo CCD. Dado que el raspberry pi no tiene la capacidad de hacer exposiciones largas con el módulo CCD y hacer otras operaciones al mismo tiempo, el resto del control es delegado a una tarjeta de desarrollo arduino nano. El raspberry manda una señal al arduino para indicarle cuándo comenzar y terminar el movimiento de la línea de retardo y el espejo rotatorio.

Se eligió para su función una tarjeta arduino nano por tres principales razones. Primeramente se necesitaba una solución que incluyera la lógica para hacer el movimiento completo de la línea de retardo (como mínimo contar y cambiar de dirección), lo que descalifica soluciones con circuitos integrados simples como osciladores, o aquellos hechos específicamente para hacer señales para PWM. En segunda, que su facilidad de uso lo hace muy útil para prototipar; en el proceso de desarrollar el dispositivo hay que hacer muchas pruebas y modificar constantemente su funcionamiento, y es sumamente fácil y rápido modificar el programa que seguirá el arduino, lo que facilita el proceso. Finalmente la tercera razón fue que su facilidad de uso y de aprendizaje lo hacen muy accesible, de manera que otros alumnos que trabajen en el laboratorio en el futuro puedan modificar fácilmente

^[5] *Virtual Network Computing* o computación virtual en red, un sistema de comunicación con el cual se puede controlar una computadora de manera remota.

^[6] Como nota para el usuario del dispositivo, el raspberry fue configurado para abrir un servidor VNC con la contraseña 15634146 al iniciarse. Al encenderlo simplemente hay que conectarlo mediante un cable ethernet con una computadora con un cliente VNC y se podrá acceder al dispositivo. También es posible configurar que esta conexión ocurra mediante wifi y así omitir el cable ethernet. Durante este trabajo el dispositivo fue usado a través del software VNC Viewer. También se usó el software WinSCP para establecer una conexión SSH con el dispositivo y transmitir documentos entre el dispositivo y la computadora, esto también se puede hacer mediante la conexión VNC dependiendo del cliente que se use. Para la conexión ssh se necesita notar que el raspberry tiene el nombre de usuario *pi* con la contraseña *frog*.

cada parte del dispositivo y seguir mejorándolo^[7]. En general la fácil modificación y fácil accesibilidad a cualquier parte del dispositivo fue una consideración importante durante todo el rediseño.

Para el control de la línea de retardo, como se mencionó en la sección 6.1.1, se usa una señal PWM, una señal digital de relativamente alta frecuencia que, oscilando rápidamente entre un voltaje bajo y uno alto, controla la corriente promedio que se enviará a la línea de retardo. Para asegurar el movimiento suave de la línea de retardo, se modificaron los límites que el arduino impone sobre su función PWM, para llegar a una señal PWM con frecuencia de aproximadamente 7800KHz. Además la señal del arduino se pasa por un transistor funcionando como interruptor, para que la corriente destinada a la línea de retardo provenga directamente de la fuente de poder, y un filtro que elimina las componentes de alta frecuencia de la señal resultante de manera que a la línea de retardo llegue una rampa de voltaje continua y suave. En la primera prueba que se discutió en la sección 6.1.1 se confirmó que la señal se suavizó lo suficiente para que la línea de retardo se moviera correctamente, no se notaron movimientos bruscos ni saltos incluso en la escala de la longitud de onda del láser usado, de 644nm.

Se diseñó un circuito para ser impreso a una placa PCB (*Printed circuit board*), se muestra en la figura 6.12.

^[7]Considérese que, por ejemplo, entre los alumnos de la facultad de ciencias (muy probables candidatos a continuar el proyecto) no es muy común que sepan usar y programar un microcontrolador PIC, mientras que muchos más habrán aprendido a usar arduino.

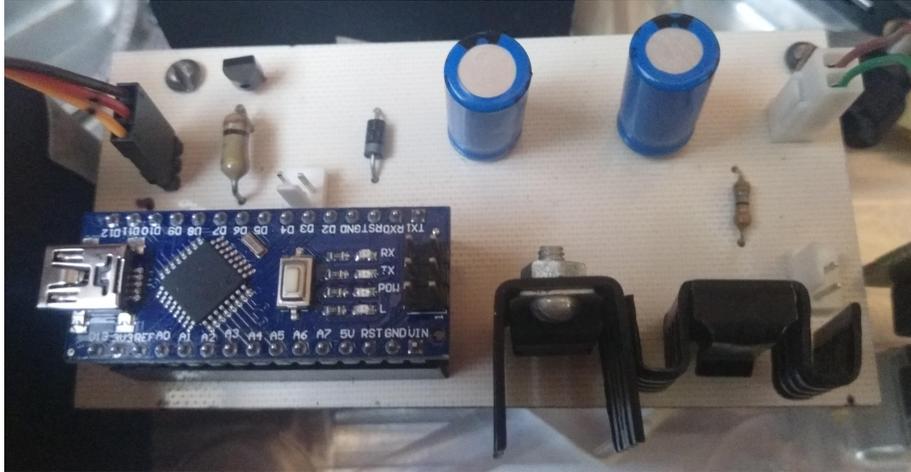


Figura 6.12: Circuito en PCB que forma parte del sistema de control.

El circuito diseñado incluye dos reguladores de voltaje: uno de 5v que se encarga de alimentar al arduino, y otro de 8v que provee la energía que usa el interruptor para mandar a la línea de retardo. También contiene el interruptor y el arduino, montado sobre pines hembra de manera que sea removible. Finalmente contiene el filtro mencionado anteriormente que suaviza la señal para la línea de retardo.

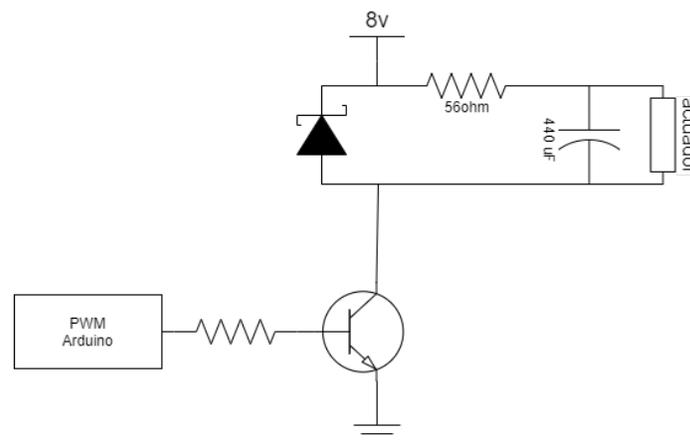


Figura 6.13: Diagrama de la parte del circuito que controla la línea de retardo, incluyendo un transistor como interruptor y el filtro que suaviza la señal.

El diseño del filtro se realizó con dos consideraciones en mente: debe tener una frecuencia de corte baja para eliminar todas las componentes de alta frecuencia que sea posible, y

debe tener una resistencia suficientemente pequeña para no limitar la corriente que puede atravesar la línea de retardo y por lo tanto limitar su rango de movimiento.

Lo ideal sería usar una resistencia muy pequeña para no limitar la corriente, sin embargo esto requeriría usar un condensador de valor muy grande. La solución a la que se llegó para no limitar el rango en el diseño final fue aumentar el voltaje introducido a esa sección del circuito de 5V a 8V, contrarrestando la limitación en corriente que causa la resistencia. La frecuencia de corte del filtro es de aproximadamente 40Hz.

Además del circuito mostrado en la figura, se añadió una sección más (actualmente se encuentra en una pequeña protoboard, en el futuro se puede modificar el diseño para ponerlo en la misma placa PCB) con un controlador tipo pololu A4988 para el control del motor a pasos que mueve el espejo rotatorio. El arduino se encarga de mandar la señal al controlador para controlar el movimiento del motor. La lógica del controlador es alimentada por el mismo regulador que alimenta el arduino. El voltaje que requiere el motor se obtiene directamente del eliminador que alimenta al circuito antes de los reguladores. El controlador tiene funcionalidad de micropasos, es decir, de dar pasos que son fracciones del paso normal del motor. En el caso del controlador A4988, se pueden dar micropasos hasta de un dieciseisavo del paso completo. Hay un controlador similar con función de micropasos de treintidosavos, con nombre DRV8825, sin embargo se encontró que los micropasos con este controlador eran inconsistentes.

El módulo CCD se controla mediante la librería de python propia del Raspberry para ese propósito. Es importante durante el control del módulo desactivar los varios ajustes automáticos que realiza por default el módulo CCD en su uso como cámara, para asegurar que las trazas obtenidas sean repetibles y que se puedan obtener correctamente, que no haya sobresaturación. Estos ajustes incluyen el ISO, el balance de blancos y, críticamente, el tiempo de exposición. El control de la cámara durante el funcionamiento del aparato consiste en comenzar una exposición larga en cuanto se mande la señal para iniciar el movimiento de la línea de retardo y del espejo rotatorio, de manera que durante el tiempo de exposición se capturen un número entero de ciclos del movimiento. El tiempo de exposición resulta ser crítico por esta razón, dado que tiene que corresponder a un número entero

de ciclos de la línea de retardo, está limitado por la velocidad de movimiento de todas las partes móviles, por lo que hay un balance al que se debe llegar para evitar imágenes sobresaturadas.

Posteriormente hay dos opciones de cómo expresar la información. La primera es con una codificación conocida como YUV^[8]. A diferencia de la codificación RGB más común, que expresa tres valores de intensidad para cada pixel, las intensidades de luz roja, verde y azul necesarias para recrear el color e intensidad de ese pixel, la codificación YUV guarda para cada pixel un valor de intensidad total, o luminancia (al que llama Y), y dos valores que describen su color (U y V)^[9]. Al usar esta codificación, se puede simplemente descartar la información de color directamente de la información que da el módulo CCD, y así generar la traza con sólo la intensidad. La segunda opción es recibir la información en formato RGB y simplemente descartar la información de los canales R (rojo) y G (verde). Esto es razonable debido a que la luz que llega al sensor es azul, alrededor de los 400nm, y tiene un ancho espectral muy pequeño, de manera que cae completamente dentro del rango de sensibilidad del canal azul. Podría parecer a primera vista que el resultado de cualquiera de estas dos opciones debería ser el mismo, sin embargo no es el caso, pues a pesar de que la luz involucrada es completamente azul, a altas intensidades o altos tiempos de exposición los sensores que corresponden a los otros canales en la CCD sí muestran una respuesta.

Así se obtiene una imagen de la traza, que puede posteriormente ser procesada para ser introducida al software de análisis. El procesamiento que se debe hacer antes de pasar al software de análisis incluye asegurar que la traza esté centrada en la imagen y reducir la resolución de la imagen a la que acepta el software. También es recomendable procesar la imagen con operaciones de filtrado para reducir el ruido, restar el nivel de intensidad de fondo que se pueda presentar, y asegurar que no haya señal en las orillas y esquinas de la imagen.

^[8]Mientras que el nombre de la codificación RGB (Red Green Blue) hace referencia a los colores que representan sus valores, rojo, verde y azul, en YUV las letras no representan palabras, son letras arbitrarias que se escogieron para los valores abstractos que representan.

^[9]La versión específica de la codificación usada por el software del raspberry, YUV420, en realidad tiene un valor de luminancia por pixel, pero un valor de cada componente de color para cada 4 pixeles, es decir, la información de intensidad tiene 4 veces la resolución de la información de color.

Capítulo 7

Pruebas al dispositivo

Como se mencionó en la descripción del dispositivo en el capítulo 6, se probaron dos motores para la sección espectrométrica del FROG. En este capítulo se describirán las pruebas que se hicieron con ambos para determinar las características que se necesitan en el diseño final. Son de principal importancia la precisión del motor, que sea capaz de hacer movimientos pequeños, la repetibilidad de esos movimientos, que pueda regresar a la misma posición después de cada oscilación, y finalmente la suavidad o continuidad del movimiento, que aseguraría una traza continua. En general los dos tipos de motor que podrían cumplir estas características son servomotores y motores a pasos. Un motor DC no tiene la capacidad de control de posición que se necesita para hacer el movimiento repetible. Si se añadiera lo que se conoce como un *codificador rotatorio*, un motor DC podría ser una opción, sin embargo esta opción necesita un sistema de control considerablemente más complejo, así que se decidió explorar las otras dos opciones primero en este proyecto.

Tanto los servomotores como los motores a pasos cumplen con el requerimiento de tener control repetible sobre su posición; los servomotores se controlan directamente con su posición, y con los motores a pasos es simple contar los pasos para saber la posición en cada momento. El requerimiento de precisión en el caso de los servomotores depende fuertemente de la calidad del servomotor usado y de el sistema que se use para controlarlo, por dar un ejemplo básico, la librería más común de arduino para controlar servomotores solo

tiene la opción de especificar la posición en números enteros de grados^[1]. Dependiendo del servomotor usado, puede ser que la precisión mecánica del motor sea mejor o peor que la del sistema de control usado. La precisión de un motor a pasos evidentemente depende sólo del tamaño del paso. Como se mencionó en el capítulo anterior se puede usar un controlador con función de micropasos para disminuir drásticamente el tamaño de los pasos, lo que potencialmente hace que sea viable el uso de motores a pasos a pesar de que inicialmente puede parecer que no lo es. También existe en ambos casos la posibilidad de usar un motorreductor para aumentar mecánicamente la precisión del movimiento. En cuanto a la suavidad del movimiento, el servomotor parece ser la opción clara, pero existe el problema de la velocidad de movimiento. El tipo más común de servomotores en electrónica controlan solo su posición, se mueven de una posición a otra a una velocidad específica. Para reducir su velocidad se tendría que usar esencialmente como un motor a pasos, especificándole posiciones intermedias en su movimiento a la velocidad que se quiera, pero usarlo así eliminaría su ventaja sobre el motor a pasos. El motor a pasos evidentemente no se mueve de manera suave, pero si se consiguiera moverlo con pasos suficientemente pequeños, sería posible aproximar un movimiento suave.

Para ambas pruebas que se describirán a continuación el procedimiento fue el siguiente: se alineó el dispositivo con la línea de retardo fija, de manera que se genere un espectro fijo, que no cambiará durante el movimiento. Posteriormente se hace mover el espejo rotatorio y se obtiene una imagen como se haría durante el uso normal del dispositivo. Dado que no cambia la forma del espectro, lo que se busca ver es que la oscilación del espejo genere en la imagen un rectángulo uniforme, resultado de barrer una línea recta en el eje perpendicular a ella. Esta prueba busca poner a prueba la suavidad del movimiento del motor, si el movimiento no es suficientemente suave y continuo, se deformará el rectángulo, y esta deformación le ocurrirá también a una traza FROG obtenida con el dispositivo.

^[1]en general los servomotores comunes se controlan con un tren de pulsos donde la longitud de los pulsos corresponde a la posición, cualquier sistema de control va a tener un límite en el incremento mínimo de duración de los pulsos.

7.1. Servo SG90

Se usó un servomotor barato, simple y accesible principalmente para investigar la posible utilidad de un servomotor pequeño para este uso sin gastar innecesariamente antes de tener una idea de cuales pueden ser las limitaciones de este tipo de motor en el diseño. Se efectuó la prueba como se describió anteriormente, se muestra en la figura 7.1 el resultado.

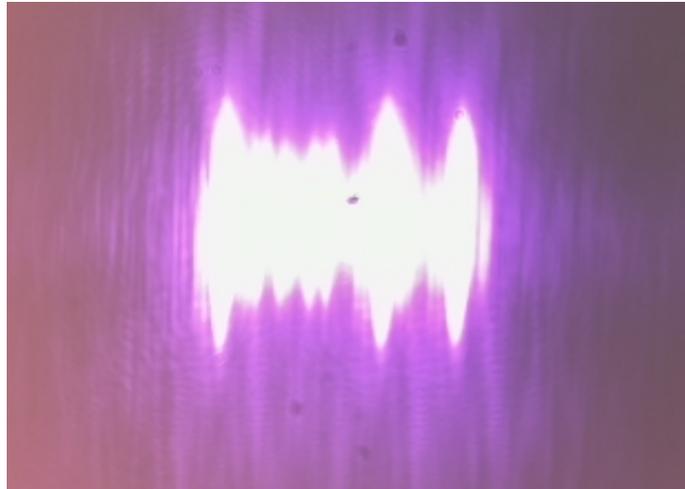


Figura 7.1: Resultado de un barrido de un espectro constante con el servomotor.

Se observan varias cosas. Es evidente que la imagen está sumamente sobresaturada, idealmente la figura no se vería totalmente blanca, como el espectro obtenido en la figura 6.11. Esta sobreexposición evidente se debe a que el movimiento del espejo es lento, y se requiere exponer la imagen por un tiempo muy largo para captar todo el rango de movimiento. Si se estuviera tomando una traza real, se estaría perdiendo información debido a la sobreexposición, sin embargo para el propósito de esta prueba no supone un problema. La sobreexposición se podría mitigar modificando los ajustes del módulo CCD, la velocidad del movimiento (aunque esta estará limitada por la línea de retardo), asegurando que el módulo CCD esté completamente aislado de la luz ambiental, y finalmente atenuando el haz antes de que alcance el módulo.

La observación más importante es que la figura que se forma no es un rectángulo uni-

forme. El cambio en el ancho de la figura a lo largo de ella se debe a que el espejo se mueve de manera discontinua e impredecible, en las secciones donde pasó más tiempo las orillas del espectro generaron más exposición en la imagen, haciendo que la figura parezca más ancha en esa sección. Era natural esperar que esto ocurriera en los extremos de la figura debido a que ahí el espejo debe desacelerar y acelerar en dirección contraria. Éste comportamiento en los extremos no necesariamente es un problema, dado que el movimiento de la línea de retardo debe incluir los extremos donde los pulsos no se empalman, y por tanto en estas secciones no debe haber luz en el uso normal del dispositivo. El hecho de que este comportamiento irregular ocurra en la parte central de la figura muestra que el movimiento de este motor no es apto para usarse en el dispositivo. En varias repeticiones de esta prueba se encontró siempre este comportamiento irregular pero de manera impredecible, es decir, no siempre se observa la misma forma, por lo que no existe manera de compensar su efecto.

Con base en este resultado se hicieron más pruebas donde se observó que el servomotor no era capaz de efectuar movimientos pequeños de manera repetible, lo que causa que no se mueva de manera suave a las escalas necesarias.

Viendo estos problemas, se puede concluir que si se quisiera usar un servomotor, debe ser capaz de controlar su velocidad de movimiento. Existen servomotores que controlan la velocidad de movimiento y no la posición, un motor de este tipo no sería útil, pues se necesita que cada vez que se quiera hacer un barrido, el espejo regrese exactamente al mismo ángulo, con una tolerancia muy pequeña. Por lo tanto, para usar un servomotor se necesitaría uno con tres características importantes: Control de posición, control de velocidad y alta precisión a ángulos pequeños de movimiento. Cabe mencionar que esta última característica casi nunca se reporta en las hojas de datos, por lo que podría ser complicado encontrar un servomotor con esta característica.

7.2. Motor a pasos Nema 17 con micropasos

En cuanto al motor a pasos se realizó la misma prueba, y además se tomaron capturas del espectro paso a paso para analizar la uniformidad de los pasos. Se muestran en las figuras 7.2 y 7.4 los resultados respectivos de las pruebas. Cabe mencionar que debido a que la montura mencionada en la sección 6.2 fue diseñada para un motor más pequeño no fue posible usarla para esta prueba. Esto complicó la alineación y resultó en un movimiento ligeramente desviado del espectro. Sin embargo este problema no afecta los aspectos de la imagen que se busca analizar, por lo que no supone un problema para las conclusiones a las que se llegará a continuación.

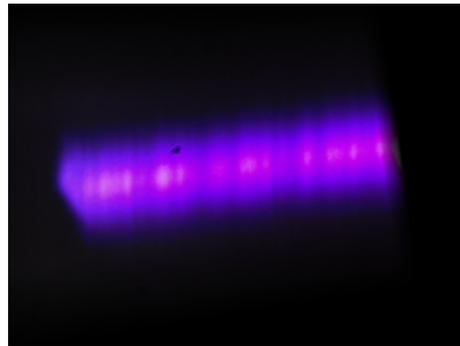


Figura 7.2: Resultado de un barrido de un espectro constante con el motor a pasos.

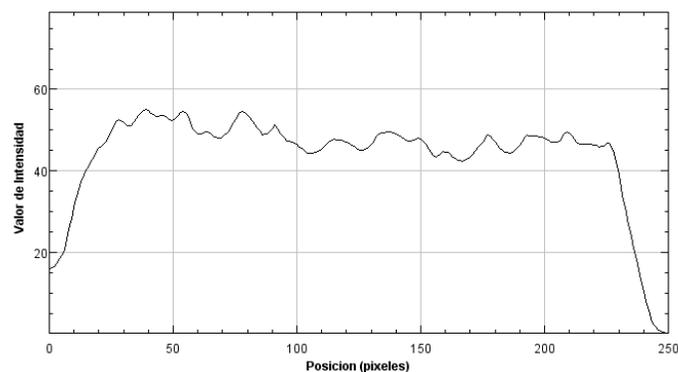


Figura 7.3: Gráfica del marginal de retardo de la imagen 7.2.

En la figura 7.2 se observa que el resultado es mucho más suave que el que se obtuvo

con el servomotor. Sin embargo, sí se aprecia una ondulación en la forma de las orillas, y también una variación en la intensidad al centro de la figura. Evidentemente esta variación se debe a que los pasos del motor son demasiado grandes, por lo que no asemejan un movimiento uniforme. La variación se puede cuantificar usando el marginal de retardo, discutido previamente en la sección 3.1. Usando el software de análisis de imágenes ImageJ se midió el valor promedio de intensidad en cada columna de la imagen (equivalente al marginal normalizado), que se muestra en la figura 7.3. Idealmente se observaría una meseta completamente plana que cae bruscamente a las orillas. Es importante notar que el comportamiento de los pasos del motor es sumamente repetible, a diferencia del comportamiento del servomotor visto en la sección anterior. Esto abre la posibilidad de usar la técnica descrita al final de la sección 3.1 pero con el marginal de retardo para compensar las variaciones y eliminar su influencia sobre la traza. Para resolver el problema mecánicamente se podría usar un motoreductor con el motor a pasos para obtener la situación que se buscaba inicialmente, donde los pasos sean suficientemente pequeños para no causar este tipo de variaciones.

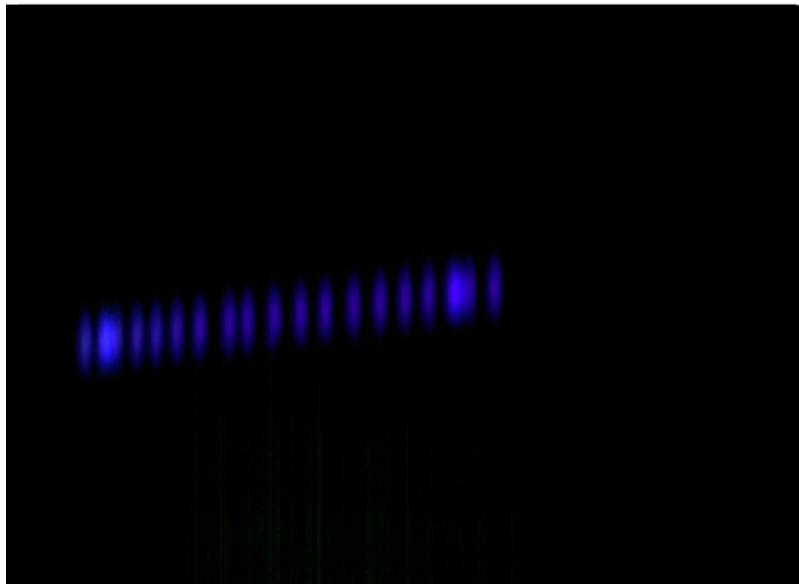
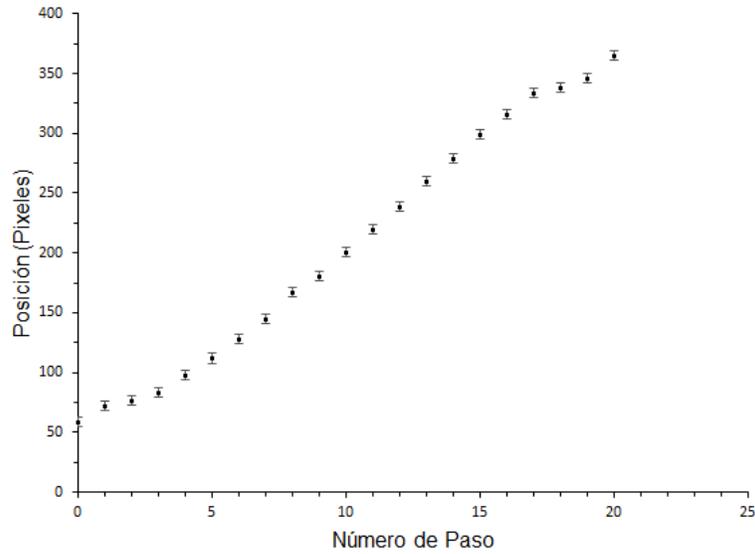


Figura 7.4: izquierda: Gráfica de la posición relativa del espectro para cada paso. Derecha: Imagen compilada con cada posición del espectro

En la figura 7.4 se muestra el espectro observado para cada micropaso del motor, compilados en una sola imagen, así como una gráfica de su posición relativa en la imagen contra cada paso. La observación más preocupante y evidente concierne a los pasos 2, 3, 4, 18, 19 y 20. Se observa que en estas regiones se empalman casi completamente las imágenes de los pasos (al punto que es difícil distinguirlos). Lo que esto nos dice es que en estas regiones el motor da pasos más pequeños. Este problema se debe a la función de micropasos

del controlador; se observa que el comportamiento problemático se repite exactamente cada 16 micropasos, es decir, cada paso entero del motor. Se observó durante las pruebas que este problema es altamente dependiente de los ajustes del controlador, específicamente de su función de limitación de voltaje, por lo que el problema se podría disminuir cambiando el voltaje que se introduce al controlador, o en el peor de los casos, cambiando completamente de controlador ^[2]. Al igual que en el párrafo anterior, este problema también se podría resolver usando un motorreductor en vez de la función de micropasos del controlador, asegurando así que cada paso sea igual pero aún suficientemente pequeño.

^[2]Cabe mencionar que este problema se observó con dos controladores diferentes, un controlador tipo Pololu A4988, y un controlador tipo Pololu DRV8825, este último siendo el más problemático de los dos.

Capítulo 8

Conclusiones

En este trabajo se describió la teoría relacionada al problema de la medición de pulsos ultracortos mediante la técnica FROG. Se dio una descripción general de los pulsos ultracortos, así como de los fenómenos de la óptica no lineal que se usan para medirlos. De particular importancia es la generación de segundo armónico, pues ésta es la que se usa en la técnica SHG-FROG. La descripción de la técnica FROG no estaría completa sin hablar del algoritmo FROG, y de cómo extrae la información de la traza FROG para obtener las características del pulso que la generó.

Se implementó un software de análisis en python para dar al usuario una herramienta sencilla y al alcance para analizar las trazas que obtenga con el aparato. Es importante recordar que se deben procesar las imágenes crudas que se obtienen de la CCD para que cumplan las características de una traza FROG como se describen en la sección 3.1. Para este fin, se pueden hacer casi todas las modificaciones necesarias con software científico de análisis de imágenes como ImageJ o similares, la parte más complicada del procedimiento es aplicarle un filtro sencillo para el ruido aleatorio, y restar el nivel de fondo. El software demostró su capacidad de recuperar pulsos de trazas simuladas, resta solo probarlo con las trazas reales que se obtengan con el dispositivo.

En su versión actual, el dispositivo SHG-FROG muestra potencial para ser capaz de ha-

cer mediciones de trazas hasta en menos de un segundo. Se caracterizó la línea de retardo, y se encontró, como se muestra en la figura 6.8 un rango de movimiento lineal del tamaño necesario para medir pulsos de hasta $300fs$, en el cual se puede oscilar a una frecuencia de alrededor de 2 Hz sin causar desalineación. Esto permitiría, dependiendo de cuantos barridos se quieran usar para cada traza, obtener trazas hasta en un segundo. Esto representa una mejora de dos órdenes de magnitud respecto al diseño anterior. El rediseño se enfocó en este aumento de velocidad, así como en accesibilidad al usuario y facilidad de modificación, de manera que cualquier usuario futuro pueda implementar los cambios que se requieran para adaptar el funcionamiento del dispositivo a las necesidades del laboratorio.

Respecto a la sección espectrométrica del aparato, se analizó el uso de dos tipos distintos de motores, sus ventajas y desventajas. En el caso de un servomotor, se concluyó que se tendría que buscar uno con especificaciones muy particulares, que sea capaz de controlar tanto posición como velocidad, y que el sistema de control podría incluso en ese caso ser un factor limitante. Otra opción sería conseguir un elemento especializado como un galvanómetro de espejo, o galvo, para cumplir esta función. Con el motor a pasos que se probó se encontraron buenos resultados, limitados principalmente por el uso de controladores que permiten micropasos. Se encontró como se muestra en las figuras 7.2 y 7.3 un movimiento relativamente suave, pero que aún no es tan uniforme como se necesita para el funcionamiento correcto del dispositivo.

Se reemplazó completamente el sistema de control, ahora consiste principalmente en una raspberry pi, con un arduino controlando la línea de retardo mediante PWM. La raspberry permite controlar el aparato remotamente y tener una interfaz con el usuario. Al momento el código que controla el aparato se encuentra como un script de python que se debe correr en la terminal del Raspberry Pi. Este script manda la señal que necesita el arduino para correr su propio código, que a su vez se puede ver y modificar fácilmente. Además controla el módulo CCD, y regresa como resultado una imagen que, si todo está en orden, debe contener la traza FROG.

Como trabajo a futuro en términos del espectrómetro, queda la decisión de qué motor se usará como espejo rotatorio. Hay dos opciones viables para terminar esta sección. La

primera es incorporar un motorreductor al motor a pasos, lo que implicaría el diseño de una nueva manera de montar y acoplar al motor y el motorreductor, similar al desarrollo de una pieza impresa en 3D que se describe en la sección 6.2. La segunda opción es conseguir un galvo e incorporarlo al diseño. Para este fin se podría reutilizar la montura diseñada en la sección 6.2, dependiendo del tamaño que tenga el galvo. El principal problema sería acoplarlo al sistema de control, dependiendo de cómo se controle. Es posible que se pueda usar la misma señal que controla la línea de retardo, pero se necesitaría adaptarla para controlar este nuevo elemento.

Para asegurar que se obtengan trazas adecuadas repetiblemente, se deberá afinar el funcionamiento del sistema de control y determinar varias configuraciones cuando sea posible comenzar a obtener trazas, por ejemplo el tiempo de exposición necesario, el número de barridos que se medirán en una sola imagen y el rango de movimiento ideal. Dada la velocidad de movimiento de la línea de retardo resultan particularmente importantes estos factores. Se observó durante las pruebas descritas en la última sección que el tiempo de exposición necesario para captar todo el movimiento de la línea de retardo a su máxima velocidad resulta muy largo, y la imagen queda sobreexpuesta con el tiempo equivalente a una sola pasada. Es probable entonces que sea necesario poner un elemento atenuador frente a la cámara para permitirle captar luz por más tiempo sin causar sobreexposición. Incluso en ese caso será clave afinar los parámetros de la cámara para obtener una buena imagen, particularmente si se quiere utilizar varias pasadas. También es importante evitar que la luz del ambiente genere ruido de fondo en la cámara, y este problema se multiplica al usar varias pasadas, pues la cámara permanece expuesta al ruido aún más tiempo.

Anexo A

Algunos cálculos

A continuación se desarrollará el cálculo de la polarización de segundo orden para el caso de dos ondas como se describen en la ecuación 2.8, reproducida a continuación.

$$\mathcal{E}(t) = E_1(t)e^{i\omega_1 t} + E_2(t)e^{i\omega_2 t} + c.c. \quad (2.8 \text{ Reproducida})$$

Para hacer los cálculos con todo detalle, expandamos los complejos conjugados y consideremos la parte espacial de las ondas, es decir, su dependencia en \mathbf{r} .

$$\mathcal{E}(\mathbf{r}, t) = E_1(\mathbf{r}, t)e^{i(\omega_1 t - \mathbf{k}_1 \cdot \mathbf{r})} + E_1^*(\mathbf{r}, t)e^{-i(\omega_1 t - \mathbf{k}_1 \cdot \mathbf{r})} + E_2(\mathbf{r}, t)e^{i(\omega_2 t - \mathbf{k}_2 \cdot \mathbf{r})} + E_2^*(\mathbf{r}, t)e^{-i(\omega_2 t - \mathbf{k}_2 \cdot \mathbf{r})} \quad (A.1)$$

Para ahorrar espacio, de aquí en adelante se escribirán $E_1(\mathbf{r}, t)$ y $E_2(\mathbf{r}, t)$ como simplemente E_1 y E_2 . La expresión A.1 debe introducirse en la ecuación 2.5 reproducida a continuación.

$$\tilde{P}^{(2)}(t) = \epsilon_0 \chi^{(2)} \mathcal{E}(t)^2 \quad (2.5 \text{ Reproducida})$$

por lo tanto, en este caso:

$$\tilde{P}^{(2)}(\mathbf{r}, t) = \epsilon_0 \chi^{(2)} \left(E_1 e^{i(\omega_1 t - \mathbf{k}_1 \cdot \mathbf{r})} + E_1^* e^{-i(\omega_1 t - \mathbf{k}_1 \cdot \mathbf{r})} + E_2 e^{i(\omega_2 t - \mathbf{k}_2 \cdot \mathbf{r})} + E_2^* e^{-i(\omega_2 t - \mathbf{k}_2 \cdot \mathbf{r})} \right)^2$$

Distribuyendo los términos y dividiendo las constantes para que no estorben:

$$\begin{aligned} \frac{\tilde{P}^{(2)}(\mathbf{r}, t)}{\epsilon_0 \chi^{(2)}} = & E_1 e^{i(\omega_1 t - \mathbf{k}_1 \cdot \mathbf{r})} \left(E_1 e^{i(\omega_1 t - \mathbf{k}_1 \cdot \mathbf{r})} + E_1^* e^{-i(\omega_1 t - \mathbf{k}_1 \cdot \mathbf{r})} + E_2 e^{i(\omega_2 t - \mathbf{k}_2 \cdot \mathbf{r})} + E_2^* e^{-i(\omega_2 t - \mathbf{k}_2 \cdot \mathbf{r})} \right) \\ & + E_1^* e^{-i(\omega_1 t - \mathbf{k}_1 \cdot \mathbf{r})} \left(E_1 e^{i(\omega_1 t - \mathbf{k}_1 \cdot \mathbf{r})} + E_1^* e^{-i(\omega_1 t - \mathbf{k}_1 \cdot \mathbf{r})} + E_2 e^{i(\omega_2 t - \mathbf{k}_2 \cdot \mathbf{r})} + E_2^* e^{-i(\omega_2 t - \mathbf{k}_2 \cdot \mathbf{r})} \right) \\ & + E_2 e^{i(\omega_2 t - \mathbf{k}_2 \cdot \mathbf{r})} \left(E_1 e^{i(\omega_1 t - \mathbf{k}_1 \cdot \mathbf{r})} + E_1^* e^{-i(\omega_1 t - \mathbf{k}_1 \cdot \mathbf{r})} + E_2 e^{i(\omega_2 t - \mathbf{k}_2 \cdot \mathbf{r})} + E_2^* e^{-i(\omega_2 t - \mathbf{k}_2 \cdot \mathbf{r})} \right) \\ & + E_2^* e^{-i(\omega_2 t - \mathbf{k}_2 \cdot \mathbf{r})} \left(E_1 e^{i(\omega_1 t - \mathbf{k}_1 \cdot \mathbf{r})} + E_1^* e^{-i(\omega_1 t - \mathbf{k}_1 \cdot \mathbf{r})} + E_2 e^{i(\omega_2 t - \mathbf{k}_2 \cdot \mathbf{r})} + E_2^* e^{-i(\omega_2 t - \mathbf{k}_2 \cdot \mathbf{r})} \right) \end{aligned}$$

Se continúa distribuyendo:

$$\begin{aligned} \frac{\tilde{P}^{(2)}(\mathbf{r}, t)}{\epsilon_0 \chi^{(2)}} = & E_1^2 e^{i(2\omega_1 t - 2\mathbf{k}_1 \cdot \mathbf{r})} + E_1 E_1^* + E_1 E_2 e^{i[(\omega_1 + \omega_2)t - (\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{r}]} + E_1 E_2^* e^{i[(\omega_1 - \omega_2)t - (\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{r}]} \\ & + E_1^* E_1 + E_1^{*2} e^{-i(2\omega_1 t - 2\mathbf{k}_1 \cdot \mathbf{r})} + E_1^* E_2 e^{i[(\omega_2 - \omega_1)t - (\mathbf{k}_2 - \mathbf{k}_1) \cdot \mathbf{r}]} + E_1^* E_2^* e^{-i[(\omega_1 + \omega_2)t - (\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{r}]} \\ & + E_1 E_2 e^{i[(\omega_1 + \omega_2)t - (\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{r}]} + E_1^* E_2 e^{i[(\omega_2 - \omega_1)t - (\mathbf{k}_2 - \mathbf{k}_1) \cdot \mathbf{r}]} + E_2^2 e^{i(2\omega_2 t - 2\mathbf{k}_2 \cdot \mathbf{r})} + E_2 E_2^* \\ & + E_1 E_2^* e^{i[(\omega_1 - \omega_2)t - (\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{r}]} + E_1^* E_2^* e^{-i[(\omega_1 + \omega_2)t - (\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{r}]} + E_2^* E_2 + E_2^{*2} e^{-i(2\omega_2 t - 2\mathbf{k}_2 \cdot \mathbf{r})} \end{aligned}$$

Juntando términos similares y acomodando:

$$\begin{aligned} \frac{\tilde{P}^{(2)}(\mathbf{r}, t)}{\epsilon_0 \chi^{(2)}} = & E_1^2 e^{i(2\omega_1 t - 2\mathbf{k}_1 \cdot \mathbf{r})} + E_1^{*2} e^{-i(2\omega_1 t - 2\mathbf{k}_1 \cdot \mathbf{r})} + E_2^2 e^{i(2\omega_2 t - 2\mathbf{k}_2 \cdot \mathbf{r})} + E_2^{*2} e^{-i(2\omega_2 t - 2\mathbf{k}_2 \cdot \mathbf{r})} \\ & + 2E_1 E_2 e^{i[(\omega_1 + \omega_2)t - (\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{r}]} + 2E_1^* E_2^* e^{-i[(\omega_1 + \omega_2)t - (\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{r}]} \\ & + 2E_1 E_2^* e^{i[(\omega_1 - \omega_2)t - (\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{r}]} + 2E_1^* E_2 e^{i[(\omega_2 - \omega_1)t - (\mathbf{k}_2 - \mathbf{k}_1) \cdot \mathbf{r}]} \\ & + 2E_1 E_1^* + 2E_2 E_2^* \end{aligned}$$

Y si se dejan los complejos conjugados indicados de nuevo como c.c.

$$\begin{aligned} \tilde{P}^{(2)}(\mathbf{r}, t) = & \epsilon_0 \chi^{(2)} \left(E_1^2 e^{i(2\omega_1 t - 2\mathbf{k}_1 \cdot \mathbf{r})} + E_2^2 e^{i(2\omega_2 t - 2\mathbf{k}_2 \cdot \mathbf{r})} + c.c. \right) \\ & + \epsilon_0 \chi^{(2)} \left(2E_1 E_2 e^{i[(\omega_1 + \omega_2)t - (\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{r}]} + 2E_1 E_2^* e^{i[(\omega_1 - \omega_2)t - (\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{r}]} + c.c. \right) \\ & + \epsilon_0 \chi^{(2)} \left(2E_1 E_1^* + 2E_2 E_2^* \right) \end{aligned} \quad (\text{A.2})$$

Si se omite la parte espacial (i.e. se considera como parte de $E_i(\mathbf{r}, t)$), se obtiene la ecuación 2.9:

$$\begin{aligned}\tilde{P}^{(2)}(t) &= \epsilon_0 \chi^{(2)} \left(E_1(t)^2 e^{i2\omega_1 t} + E_2(t)^2 e^{i2\omega_2 t} + c.c. \right) \\ &+ \epsilon_0 \chi^{(2)} \left(2E_1(t)E_2(t)e^{i(\omega_1+\omega_2)t} + 2E_1(t)E_2^*(t)e^{i(\omega_1-\omega_2)t} + c.c. \right) \\ &+ 2\epsilon_0 \chi^{(2)} \left(E_1(t)E_1^*(t) + E_2(t)E_2^*(t) \right) \quad (2.9 \text{ Reproducida})\end{aligned}$$

Anexo B

Información práctica sobre el software de análisis

Como se mencionó en el capítulo correspondiente, el script de análisis está comentado extensivamente para ser entendido de manera fácil. Sin embargo, dado que es un script sin interfaz de usuario, se necesita cierto conocimiento básico de programación para usarlo. Este anexo tiene el propósito de servir como documentación extra para el script y facilitar su uso. Lo primero que es prudente aclarar es que el script es una implementación relativamente básica del algoritmo de análisis, hay que considerar que fue hecho por un estudiante de licenciatura y no un programador profesional, y existen por supuesto softwares comerciales e incluso algunas implementaciones compartidas en la comunidad para plataformas como matlab que pueden ser más eficientes y efectivas. Los softwares comerciales usan varios algoritmos e incluso intercambian entre ellos dinámicamente para optimizar el proceso. El propósito de el script que se escribió en este trabajo es simplemente proveer al usuario del dispositivo FROG una opción conveniente, rápida, transparente (es decir, que no sea una caja negra sin explicación de como funciona) y flexible de analizar las trazas que obtenga, no competir con o proveer mejoras en términos de rendimiento sobre otras opciones que puedan estar disponibles.

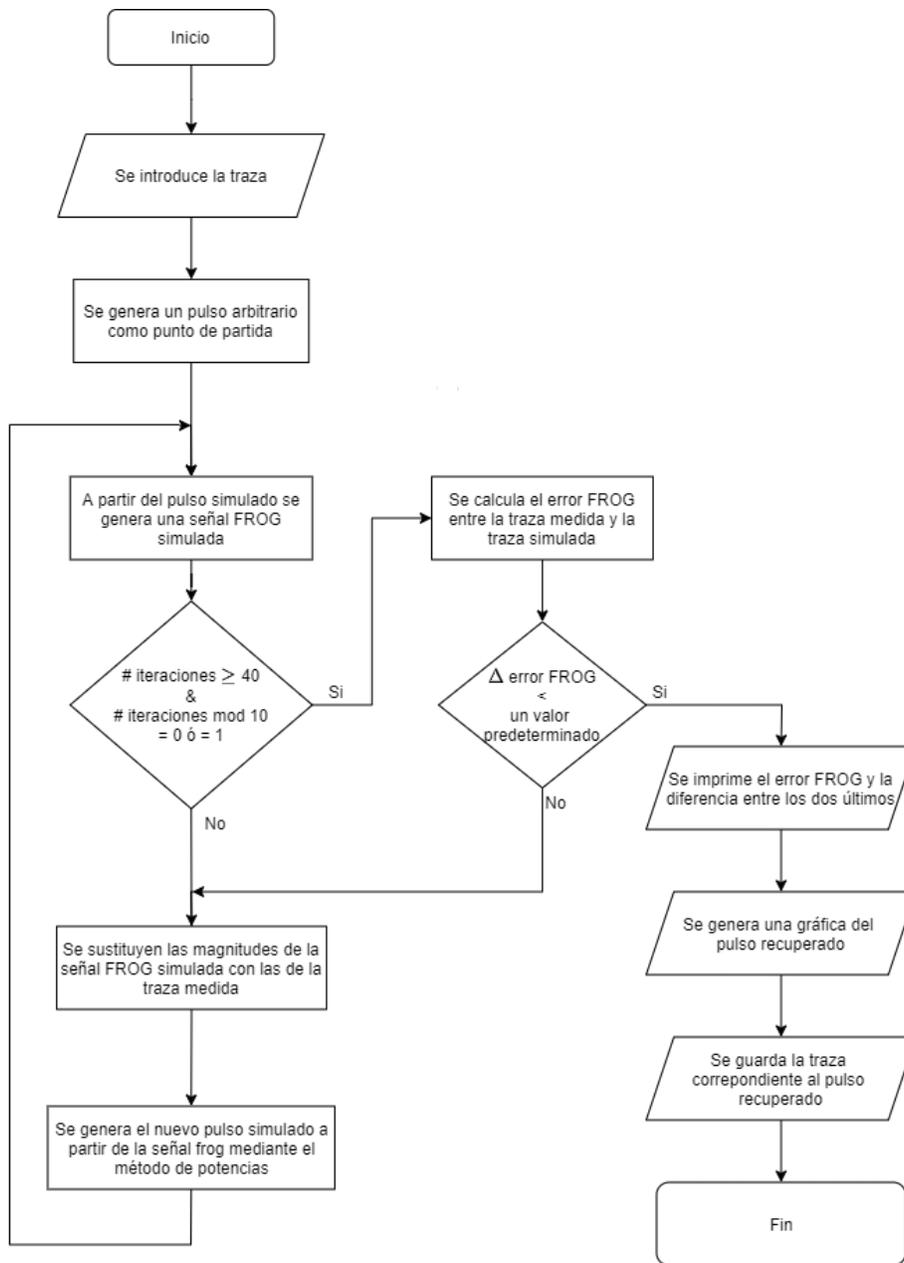


Figura 5.9: Diagrama de flujo que representa el funcionamiento del script de análisis. Reproducida de la página 48

El script se encontrará, para que no se pierda y siempre esté disponible a los estudiantes que lleguen a usar el dispositivo, dentro de la memoria del Raspberry. Sin embargo no debe ser ejecutado por el raspberry, no tiene el poder de procesamiento necesario, y por tanto no es práctico usarlo para tal propósito. La idea es extraer este script, y las trazas obtenidas por el dispositivo, para hacer el análisis en una computadora. Refiérase al Anexo B para más información sobre cómo hacer esto.

En el script se usan las librerías siguientes: la librería de manejo de imágenes PIL, numpy, el módulo optimize de la librería scipy, y el módulo pyplot de la librería matplotlib. Todas estas son librerías comunes de python y son fáciles de conseguir. Como se mencionó en el capítulo 5, la traza a analizar se debe introducir como una imagen monocromática cuadrada, preferiblemente donde el número de píxeles de cada lado sea una potencia de 2 ^[1]. En principio el tamaño puede ser arbitrario, pero durante este trabajo se usaron imágenes de 128 x 128 y 256 x 256, este tamaño es suficiente para recuperar cualquier pulso razonablemente sencillo. Usar trazas de mayor tamaño evidentemente generará un aumento significativo en el tiempo de ejecución del script. La imagen puede ser cualquier formato de imagen que acepte la librería PIL. Para establecer la imagen de entrada se deberá abrir el script en un editor ^[2] y modificar el campo claramente marcado con el nombre del archivo que contenga la traza, que debe estar guardado en la misma carpeta que el script, y también modificar la variable que especifica el tamaño de la imagen. Se recomienda ejecutar el script desde la terminal, ya que imprime información sobre el error FROG a la terminal, lo cual es útil para hacer pruebas y modificaciones. También se puede importar el script como una librería de python, de manera que se puede incorporar a un script que el usuario desee como parte de un análisis más amplio. Ésto también permite mas flexibilidad sobre como se usa, por ejemplo permite usar la sección del script que genera las gráficas para graficar resultados de análisis anteriores para compararlos, o modificar al vuelo los parámetros del análisis.

El “pulso” arbitrario que se genera como punto de partida para el proceso iterativo puede afectar el desempeño del proceso. Actualmente se usa como amplitud una función

^[1]Esto agiliza el algoritmo que hace las transformadas discretas de Fourier

^[2]En el proceso de este trabajo se usó el editor de texto Sublime Text, para evitarse sufrimiento el usuario puede usar el IDE o editor que le plazca

gaussiana de ancho arbitrario y como fase una serie de puntos generados al azar entre 0 y $\frac{\pi}{2}$. La generación azarosa de la fase hace que el proceso completo no esté completamente determinado, lo que puede causar que el análisis falle ocasionalmente. Sin embargo también es una fuerza del algoritmo, pues si en un caso cae en un mínimo local y no llega a la convergencia, se puede simplemente correr el proceso desde el inicio. Si fuera completamente determinado, no habría manera de arreglar este problema. Se ha encontrado que la elección del rango en el que se varía la fase puede afectar la probabilidad de convergencia.

Al final del capítulo 4 se mencionó que en el caso del FROG SHG se suele simetrizar el producto externo que se hace al inicio de cada iteración para obtener la matriz señal simulada. Durante las pruebas se encontró que en algunos casos mantener el pulso y la compuerta independientes causaba que el proceso convergiera más rápidamente y en otros casos ocurría lo contrario. Puede ser útil entonces en ciertos casos experimentar eliminando la simetrización. Esto se hace simplemente comentando el segundo término de una suma, específicamente en la línea 145. Hay un comentario en el script que indica dónde y cómo hacerlo.

Entre las líneas 165 y 200 se encuentra la sección donde se calcula el error FROG y se determina la condición de salida del ciclo. Como se mencionó anteriormente hay dos posibilidades para determinar si salir del ciclo, simplemente establecer un número de iteraciones totales, o calcular cada 10 iteraciones el cambio en el error en dos iteraciones seguidas y establecer un tamaño en el cual considerar que ya ha convergido el proceso. Este procedimiento es el que se muestra en la figura 5.9. Para usar este último se deben hacer pruebas para establecer un límite razonable. También se podría establecer un valor para el error FROG que se considere aceptablemente pequeño y, con la adición de unas dos líneas de código, usar eso como condición de salida.

Como ya se mencionó brevemente en la sección 6.3, es recomendable hacer algo de procesamiento antes de introducir las trazas al software, y de hecho es estrictamente necesario hacer ciertas modificaciones. Para empezar, es necesario que el centro de la traza coincida con el centro de la imagen, es imposible recuperar el pulso de una traza no centrada. También es necesario reescalar la imagen. Al hacer esto es muy importante considerar la

calibración, habrá una calibración que determina en la imagen que da el aparato cuál es la equivalencia entre los píxeles en cada dimensión de la imagen y los intervalos de tiempo y de frecuencia que representan. El reescalar la imagen modifica esta equivalencia, y es necesario calcular cuál es la nueva equivalencia para obtener un resultado. Una vez que se han hecho estas modificaciones, es recomendable eliminar el ruido mediante filtros y restando el nivel de luminosidad de fondo, la traza debe ser una isla sobre un mar completamente negro, de otra forma no se podrá recuperar un pulso efectivamente. Esto es especialmente importante en las orillas y esquinas de la imagen.

Bibliografía

- [Boyd, 2008] Boyd, R. W. (2008). *Nonlinear Optics*. Academic Press.
- [Castro-Marín et al., 2015] Castro-Marín, P., Kapellmann-Zafra, G., Garduño-Mejía, J., Rosete-Aguilar, M., and Román-Moreno, C. (2015). Webcam autofocus mechanism used as a delay line for the characterization of femtosecond pulses. *Review of scientific instruments*, 86.
- [DeLong et al., 1994] DeLong, K. W., Fittinghoff, D. N., Trebino, R., Kohler, B., and Wilson, K. (1994). Pulse retrieval in frequency-resolved optical gating based on the method of generalized projections. *Optics letters*, 19(24):2152–2154.
- [Diels and Rudolph, 2011] Diels, J. C. and Rudolph, W. (2011). *Ultrashort Laser Pulse Phenomena: Fundamentals, Techniques and Applications on a Femtosecond Time Scale*. Academic Press.
- [Eimerl et al., 1987] Eimerl, D., Davis, L., Velsko, S., Graham, E. K., and Zalkin, A. (1987). Optical, mechanical, and thermal properties of barium borate. *Journal of Applied Physics*, 62(5):1968–1983.
- [EKSMA-Optics, sf] EKSMA-Optics (s.f.). Thin bbo crystals for shg of ti:sapphire laser wavelength | eksma optics. <https://eksmaoptics.com/femtoline-components/femtoline-nonlinear-laser-crystals/thin-bbo-crystals-for-shg-of-ti-sapphire-laser-wavelength/>. Consultado: 17-02-2022.
- [Fienup, 1982] Fienup, J. R. (1982). Phase retrieval algorithms: a comparison. *Applied Optics*, 21.

- [Friedberg et al., 1982] Friedberg, S. H., Insel, A. J., and Spence, L. E. (1982). *Álgebra Lineal*. Publicaciones Cultural S.A.
- [Garduño-Mejía et al., 2003] Garduño-Mejía, J., Ramsay, E., Greenaway, A., and Reid, D. (2003). Real time femtosecond optical pulse measurement using a video-rate frequency-resolved optical gating system. *Review of scientific instruments*, 74.
- [Kane, 1998] Kane, D. J. (1998). Real-time measurement of ultrashort laser pulses using principal component generalized projections. *IEEE Journal of Selected Topics in Quantum Electronics*, 4.
- [Kane, 2008] Kane, D. J. (2008). Principal components generalized projections: a review. *Journal of the Optical Society of America*, 25.
- [Kapellmann, 2012] Kapellmann, G. (2012). Construcción de un sistema tipo frog portátil para la caracterización de pulsos Ópticos de femtosegundos. Master's thesis, Universidad Nacional Autónoma de México.
- [Kärtner et al., 2004] Kärtner, F. X., Morgner, U., Schibli, T., Ell, R., Haus, H. A., Fujimoto, J. G., and Ippen, E. P. (2004). *Few-Cycle Pulses Directly from a Laser*, pages 73–136. Springer Berlin Heidelberg.
- [Lemoff and Barty, 1992] Lemoff, B. E. and Barty, C. P. J. (1992). Generation of high-peak-power 20-fs pulses from a regeneratively initiated, self-mode-locked ti:sapphire laser. *Opt. Lett.*, 17(19):1367–1369.
- [Maine et al., 1988] Maine, P., Strickland, D., Bado, P., Pessot, M., and Mourou, G. (1988). Generation of ultrahigh peak power pulses by chirped pulse amplification. *IEEE Journal of Quantum Electronics*, 24(2):398–403.
- [RP-Photonics, sf] RP-Photonics ((s.f.)). Rp photonics encyclopedia - ultrashort pulses. https://www.rp-photonics.com/ultrashort_pulses.html. Consultado: 15-02-2020.
- [Schott-AG, 2017] Schott-AG (2017). Schott optical glass data sheets. https://refractiveindex.info/download/data/2017/schott_2017-01-20.pdf. Consultado: 17-02-2022.

- [Schultze et al., 2007] Schultze, M., Goulielmakis, E., Uiberacker, M., Hofstetter, M., Kim, J., Kim, D., Krausz, F., and Kleineberg, U. (2007). Powerful 170-attosecond XUV pulses generated with few-cycle laser pulses and broadband multilayer optics. *New Journal of Physics*, 9(7):243–243.
- [Sutherland, 2003] Sutherland, R. L. (2003). *Handbook of Nonlinear Optics*. Marcel Dekker.
- [Trebino, 2000] Trebino, R. (2000). *Frequency-Resolved Optical Gating: The Measurement of Ultrashort Laser Pulses*. Springer Science+Business Media.
- [Trebino and Kane, 1993] Trebino, R. and Kane, D. J. (1993). Using phase retrieval to measure the intensity and phase of ultrashort pulses: frequency-resolved optical gating. *Journal of the Optical Society of America*, 10(5).
- [Weiner, 2009] Weiner, A. M. (2009). *Ultrafast Optics*. John Wiley and Sons.

Imágenes Externas

Figura 1.1: Modificada de “Chirped-ultrashort-pulse.png”, obtenida de <https://commons.wikimedia.org/wiki/File:Chirped-ultrashort-pulse.png> bajo la licencia Creative Commons. Creada originalmente por el usuario Pgabolde de English Wikipedia.

Figura 2.1: Modificada de “Voltage Clipping.svg”, obtenida de https://commons.wikimedia.org/wiki/File:Voltage_Clipping.svg bajo la licencia Creative Commons. Creada originalmente por el usuario Jack1993jack de English Wikipedia.

Figura 3.1: Modificada de “Comparison convolution correlation.svg”, obtenida de https://commons.wikimedia.org/wiki/File:Comparison_convolution_correlation.svg bajo la licencia Creative Commons. Creada originalmente por el usuario Cmglee de English Wikipedia.