



**UNIVERSIDAD MICHOACANA DE SAN NICOLÁS  
DE HIDALGO**

**FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS**

”Mat. Luis Manuel Rivera Gutiérrez ”

---

*”ANÁLISIS DEL ESPECTRO OBTENIDO CON UN DISPOSITIVO  
MÓVIL DE UNA MUESTRA LÍQUIDA USANDO UNA RED  
NEURONAL”*

---

Tesis para obtener el título de:  
LICENCIADO EN CIENCIAS FÍSICO MATEMÁTICAS

PRESENTA:  
**Gonzalo Rafael Vargas Castañeda**

Asesor:

Doctor en Ciencias con especialidad en óptica  
Mauricio Ortiz Gutiérrez

Co-Asesor:

Doctor en Ciencias  
José Antonio González Cervera

Facultad de Ciencias Físico Matemáticas  
”Mat. Luis Manuel Rivera Gutiérrez ”  
U.M.S.N.H

Morelia, Michoacan - Agosto de 2016



*Dedicado a mi hija, mis padres, mi esposa  
y mis hermanos, gracias por siempre estar ahí.*



# Agradecimientos

Agradezco plenamente la confianza que me han dado mis padres a lo largo de toda mi vida, este trabajo de tesis ha sido una gran meta cumplida en mi vida y no cesan mis ganas de dar gracias a ustedes por la gran oportunidad que me dieron, por la vida misma y sobre todo por su gran amor incondicional que tienen sobre toda nuestra familia. Me siento infinitamente endeudado con ustedes y le doy gracias a Dios por tener unos padres tan maravillosos como lo son ustedes, gracias papá y mamá.

La gran inspiración que han sido mi hija y esposa, solo es comparable con el gran amor que les tengo a ambas. La alegría que me transmite el carisma de Ailyn hace que cualquier emoción de tristeza o angustia por un problema que tenga sea superado tan solo con verla sonreír. A lo largo de su corta vida me ha dado fuerza para lograr mis metas y sin saberlo es mi razón de seguir creciendo como humano, padre y profesionalista para ofrecerle la mejor calidad de vida que un padre pueda darle a su hija, Ailyn eres mi motivo de vivir.

Si me preguntaran si es cierto que el apoyo en un matrimonio es incondicional, contestaría que una prueba de ello es mi esposa Wendy. Le doy gracias a Dios por todo el tiempo que hemos tenido juntos; los problemas y las alegrías serán parte de nuestra vida, sin embargo el apoyo, paciencia, perdón y amor han sido y serán la base de todos los logros que hagamos juntos, gracias por todo lo que me has dado.

La unión por el amor entre hermanos no se compara con ningún lazo de amistad, por eso mis hermanos Andy y Patylu les doy gracias por estar conmigo y escucharme. Los tres somos el dolor de cabeza de nuestros papás y aunque somos totalmente distintos siempre hemos sido unidos, nos queremos a nuestra manera, nos entendemos y cuando realmente nos hemos necesitado siempre estuvimos y estaremos juntos.

Le agradezco la confianza, paciencia y todo su tiempo que dedicaron para guiarme a lo largo de este trabajo de tesis a mis asesores, el Doctor Mauricio

## IV

Ortiz Gutiérrez y el Doctor José Antonio Gonzalez Cervera, sin duda alguna son los mejores asesores que alguien de la facultad puede tener.

A mi primo el Gordo, más que mi primo, es mi compañero de aventuras, ocurrencias y grandes momentos que, desde pequeños, hicimos juntos y si bien la distancia y el tiempo nos separan, sabemos todo lo que nos une y nunca nos separará, Gonza gracias por toda una vida juntos.

Mi familia es basta y generalmente unida, si bien doy gracias a Dios por todos mis primos y tíos que tengo, ocuparía el equivalente a toda la tesis por mencionarlos. En la familia de mi papá les doy gracias por su cariño a todos mis tíos empezando por mi tío Alejandro, Marthita, Victor, Laura, Lilia, Sol, Lupita y Miriam, agradezco especialmente a toda la familia de mis tías Sol, Lupita y Miriam pues con ellos crecí y me dieron todo su cariño.

Por parte de la familia de mi mamá, están mis tíos Rene, Tere, Chuchín y Lety. Cada uno de ellos fue parte de mi infancia. Agradezco en especial a mi tía Tere pues fue ella quien me dio alojamiento y apoyo cuando llegue a Morelia, junto con su esposo mi tío Gollo en todo este tiempo han estado conmigo, muchas gracias.

Agradezco con todo mi corazón a mis abuelitos Jesús y Martha así como a Mamá Toñita y Papá Andres (Q.E.P.D.) por su amor de abuelitos, que no se puede describir con palabras, por todo su apoyo y consejos que me dieron los llevaré siempre conmigo, doy gracias a Dios por ustedes.

Le doy gracias a todos mis maestros que fueron parte de mi formación académica. Empezando por todos los maestros del CECyTEM 05 hasta mis profesores de licenciatura en la Facultad de Ciencias Fisico-Matemáticas. En especial agradezco a la profesora Erendira que me hizo viajar por toda la República y a los profesores Leyva y Lara que me apoyaron en toda cuestión en la Prepa. A mis maestros de la licenciatura Hector Pérez, Mendoza, Siddharta, Mota, Mariano Hernandez y Vega por su motivación que me transmitieron les doy gracias.

En el laboratorio de óptica estuvieron mis compañeras Alejandra y Jennifer, gracias por ayudarme en cuestiones técnicas así como al encargado del laboratorio Marco que sin él aún seguiríamos haciendo alguna alineación.

Dentro del aula me sobran los dedos del cuerpo para contar a todos los que me ayudaron con su amistad y compañerismo, pero en especial agradezco por todos los buenos y cansados ratos que viví con Jorge, Angel, Geo, Richy coreano, Paco, Venecia e Izza, gracias por su amistad.

Si mis amigos me definieran, serían mis amigos campeones y degenerados Luis, Richy y Yonhy los que más pueden definir mi estado mental actual. Desde el inicio juntos y aún hasta ahora con o sin problemas siempre como hermanos, que buen trío de cuatro hacemos.

A mis amigos de toda la vida Cash y July, desde la infancia con Cash y la casi pubertad con July, más que amigos son mi familia, mis hermanos. Gracias por su amistad.

A todas las personas que fueron parte de mi vida, que definieron quien soy ahora, gracias por ser parte de mi historia.



# Resumen

Los espectrofotómetros son instrumentos que miden parámetros de una muestra como función de la longitud de onda y están integrados por una fuente de luz de amplio espectro, un monocromador, un depósito para muestras y un detector. Estos instrumentos son de amplio uso en laboratorios de química, física de materiales y medicina entre otros. En esta tesis se presenta el diseño de un espectrofotómetro en el rango visible del espectro electromagnético y esta compuesto por un LED de luz blanca, una rejilla holográfica y la cámara de un dispositivo móvil como detector; el espectro que se genera al colocar una muestra líquida en el espectrofotómetro diseñado es analizado mediante una red neuronal. Las redes neuronales son un campo muy importante dentro de la inteligencia artificial y se inspiran en el comportamiento conocido del cerebro humano, principalmente el referido a las neuronas y sus conexiones, trata de crear modelos artificiales que solucionen problemas difíciles de resolver, tales como problemas de estabilidad en sistemas dinámicos, reconocimiento de patrones, etc., mediante técnicas algorítmicas convencionales. En esta tesis se implementó una red neuronal para analizar el espectro de una muestra y se obtuvo una eficiencia superior al 80 % para la identificación de la muestra.

**Palabras clave:** Espectrofotómetro, LED, rejilla holográfica, dispositivo móvil, red neuronal.



# Abstract

Spectrophotometers are instruments that measure parameters of a sample as a function of wavelength and are composed of a source of broad spectrum light, a monochromator, a deposit for samples and a detector. These instruments are widely used in chemistry, physics materials and medicine labs among others. The design of a spectrophotometer in the visible range of the electromagnetic spectrum is presented in this thesis and is composed of a white LED, a holographic grating and the camera of a mobile device as a detector; the spectrum generated by placing a liquid sample in the spectrophotometer designed is analyzed by a neural network. Neural networks are a very important field in artificial intelligence, and are inspired by the known behavior of the human brain, mainly referring to the neurons and their connections, try to create artificial models that solve difficult problems, such as problems stability dynamical systems, pattern recognition, etc., using conventional algorithmic techniques. In this thesis a neural network was implemented to analyze the spectrum of a liquid sample and an efficiency of 80 % was obtained for sample identification.

**Keywords:** Spectrophotometer, LED, holographic grating, mobile device, neural network.



# Índice general

<b>Agradecimientos</b>	<b>III</b>
<b>Resumen</b>	<b>VII</b>
<b>Abstract</b>	<b>IX</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Objetivos . . . . .	6
1.1.1. Objetivo general . . . . .	6
1.1.2. Objetivos específicos . . . . .	6
1.2. Descripción de capítulos . . . . .	7
<b>2. Interferencia</b>	<b>9</b>
2.1. Ecuaciones de Maxwell . . . . .	9
2.2. Interferencia . . . . .	12
2.3. Interferómetros de dos haces . . . . .	17
<b>3. Difracción y rejillas holográficas</b>	<b>23</b>
3.1. Principio de Huygens-Fresnel . . . . .	24
3.2. Difracción de Fraunhofer y Fresnel . . . . .	26
3.3. Rejillas de difracción . . . . .	28
3.3.1. Ecuación de la rejilla . . . . .	29
3.3.2. Órdenes de difracción . . . . .	31
<b>4. Espectroscopía</b>	<b>35</b>
4.1. Espectrometría Ultravioleta Visible (UV-Vis) . . . . .	37
4.2. Ley de Beer-Lambert . . . . .	38
4.3. El espectrofotómetro UV-Vis . . . . .	39
4.3.1. Partes básicas de un espectrofotómetro . . . . .	40
4.3.2. Espectrofotómetro convencional . . . . .	41

<b>5. Redes neuronales</b>	<b>43</b>
5.1. Redes neuronales biológicas . . . . .	44
5.2. Redes neuronales artificiales . . . . .	45
5.3. Función de error y formación de las redes . . . . .	48
5.4. Perceptrón multicapas . . . . .	50
5.5. Método de entrenamiento <i>Backpropagation</i> . . . . .	52
5.5.1. Notación . . . . .	54
5.5.2. Descenso de gradiente sobre el error . . . . .	54
5.6. Construcción del algoritmo . . . . .	56
<b>6. Desarrollo experimental</b>	<b>61</b>
6.1. Construcción de la rejilla . . . . .	61
6.2. Arreglo experimental . . . . .	64
6.3. Construcción de la red neuronal . . . . .	67
<b>7. Conclusiones</b>	<b>83</b>
<b>A. Algoritmo de la función: actualizarPesos</b>	<b>85</b>
<b>Apéndice</b>	<b>85</b>
<b>B. Algoritmo de la función: propagacionAdelante</b>	<b>87</b>
<b>C. Algoritmo de la función: propagacionAtras</b>	<b>89</b>
<b>D. Algoritmo de la función: AnalisisRedFinal</b>	<b>91</b>
<b>Referencias</b>	<b>95</b>

# Capítulo 1

## Introducción

Se puede hablar de la óptica de acuerdo con la convención de la *Optical Society of America*, para la cual es el estudio de la luz, de la manera como es emitida por los cuerpos luminosos, de la forma en la que se propaga a través de los medios transparentes y de la forma en que es absorbida por otros cuerpos [1]. La óptica, al estudiar los cuerpos luminosos, considera los mecanismos atómicos y moleculares que originan la luz. Al estudiar su propagación, lógicamente estudia los fenómenos luminosos relacionados con ella, como la reflexión, la refracción, la interferencia y la difracción. Finalmente, la absorción de la luz ocurre cuando la luz llega a su destino, produciendo ahí un efecto físico o químico, por ejemplo, en la retina de un ojo, en una película fotográfica, en una cámara de televisión, o en cualquier otro detector.

El área de la óptica comprende por sí misma una amplia gama de áreas de la ciencia como lo es la óptica física, comunicaciones ópticas, metrología, instrumentación óptica, etc. Una de las áreas más importantes es el análisis de materiales, en el desarrollo de la tesis se discutirá la identificación de la concentración de algún material utilizando su espectro.

Los fenómenos ópticos que ocurren en los materiales involucran interacciones entre la radiación electromagnética y los átomos, los iones y/o los electrones. Cuando la luz pasa de un medio a otro pueden ocurrir varios fenómenos: parte de la luz se transmite por el medio, parte se absorbe y/o parte se refleja. En esta tesis se tomará más importancia a la absorción que pueda tener un material.

En física, la absorción de la radiación electromagnética, como la luz, es el proceso por el cual dicha radiación es absorbida por la materia. Cuando la absorción se produce dentro del rango de la luz visible, recibe el nombre de absorción óptica. Esta radiación, al ser absorbida, puede, bien ser reemitida o bien transformarse en otro tipo de energía, como calor o energía eléctrica.

En general, todos los materiales absorben en algún rango de frecuencias. Aquellos que absorben en todo el rango de la luz visible son llamados materiales opacos, mientras que si dejan pasar dicho rango de frecuencias se les llama transparentes. Es precisamente este proceso de absorción y posterior reemisión de la luz visible lo que da color a la materia. De la cualidad de los materiales de absorber o emitir la luz que incide sobre ellos nace el concepto de espectroscopía [2].

La luz visible es físicamente idéntica a todas las radiaciones electromagnéticas. Es visible para nosotros porque nuestros ojos detectan esta estrecha banda de radiación del espectro electromagnético completo. Esta banda es la radiación dominante que emite el Sol.

Desde la antigüedad, científicos y filósofos han especulado sobre la naturaleza de la luz. La comprensión moderna de la luz comenzó con el experimento del prisma de Isaac Newton, con el que comprobó que cualquier haz incidente de luz blanca, no necesariamente procedente del Sol, se descompone en el espectro

del arco iris (del rojo al violeta). Newton tuvo que esforzarse en demostrar que los colores no eran introducidos por el prisma, sino que realmente eran los constituyentes de la luz blanca [3]. Posteriormente, se pudo comprobar que cada color corresponda a un único intervalo de frecuencias o longitudes de onda.

La espectroscopía es el estudio de la interacción entre la radiación electromagnética y la materia, con absorción o emisión de energía radiante [4]. Tiene aplicaciones en astronomía, física, química y biología, entre otras disciplinas científicas. El análisis espectral se basa en detectar la absorción o emisión de radiación electromagnética a ciertas longitudes de onda y se relacionan con los niveles de energía implicados en una transición cuántica.

Los campos de estudio de la espectroscopía son la espectroscopía de impedancia electrodinámica, espectroscopía astronómica, espectroscopía de absorción atómica, espectroscopía de rayos X, espectroscopía infrarroja, espectroscopía ultravioleta-visible, entre otras. Como se trabajará con la parte visible del espectro electromagnético se analizará más a fondo la espectroscopía ultravioleta-visible (UV-Vis).

La espectroscopía ultravioleta-visible es una espectroscopía de emisión de fotones y una espectrofotometría. Utiliza radiación electromagnética (luz) de las regiones visible, ultravioleta cercana (UV) e infrarroja cercana (IR) del espectro electromagnético, es decir, una longitud de onda entre 380 nm y 800 nm. La radiación absorbida por las moléculas desde esta región del espectro provoca transiciones electrónicas que pueden ser cuantificadas. La espectroscopía UV-visible se utiliza para identificar algunos grupos funcionales de moléculas, y además, para determinar el contenido y fuerza de una sustancia. Se utiliza de manera general en la determinación cuantitativa de los componentes de soluciones de iones de metales de transición y compuestos orgánicos altamente conjugados [2]. La manera de medir algún componen-

te en una solución es por medio de un espectrofotómetro.

El espectrofotómetro es un instrumento que permite comparar la radiación absorbida o transmitida por una solución que contiene una cantidad desconocida de soluto, y una que contiene una cantidad conocida de la misma sustancia. Esta espectrofotometría utiliza radiaciones del campo UV de 80 a 400 nm, principalmente de 200 a 400 nm (UV cercano) y de luz visible de 400 a 800 nm, por lo que es de gran utilidad para caracterizar las soluciones en la región ultravioleta-visible del espectro. Se rige por una ley muy importante: la ecuación de Beer-Lambert [2].

La principal desventaja de un espectrofotómetro es su precio, por no mencionar sus dimensiones y por ende resulta impráctico moverlo, pues todos los espectrofotómetros comerciales están por encima de los \$20000.<sup>00</sup>. Esto causa la principal motivación de esta tesis, si algo tan simple como la cámara de cualquier dispositivo móvil puede desempeñar la función del detector del espectro de una solución a analizar, bastaría con un arreglo óptico para generar dicho espectro de la solución y de un algoritmo que sea capaz de asociar el espectro analizado con su concentración específica para hacer un espectrofotómetro de bajo costo y versátil.

Para finales del 2015 alrededor de 1.91 billones de celulares inteligentes (Smartphones) son usados en todo el mundo. Y se estima que para el 2016 sean 2.16 billones de usuarios. El poder del Smartphone que abre las puertas para que sea un gran prospecto en el futuro de las herramientas médicas, tanto en laboratorios como en uso personal, radica en cuatro cualidades que son: Procesador poderoso, ubicuidad, conectividad y programabilidad [5].

Otro elemento importante en los espectrofotómetros es la rejilla de difracción que es una herramienta preferida para la se-

paración de los colores de la luz incidente. Las variedades de las rejillas de difracción son rejillas de difracción echelle, rejillas regladas, rejillas holográficas, entre otras. La rejilla holográfica es una elección ideal para experimentos de espectroscopía, ya que las rejillas holográficas exhiben una difracción significativamente más clara que la mayoría de otras rejillas al reducir la cantidad de luz extraviada. Esto resulta en una habilidad incrementada de ver líneas de absorción y emisión, con las rejillas de frecuencias espaciales mayores (1000 líneas/mm) [6]. Una rejilla de difracción holográfica ayudará para obtener el espectro de la sustancia en cuestión y una forma muy poderosa de interpretar estos espectros es por medio de una red neuronal, por sus propiedades descritas a continuación.

Se define una Red Neuronal Artificial (RNA) como modelos matemáticos inspirados en sistemas biológicos, adaptados y simulados en computadoras convencionales [7]. Las RNAs están inspiradas en el sistema biológico natural. Como es conocido, en este sistema la neurona es la unidad de procesamiento, y aunque las RNAs sean mucho menos complejas que una red neuronal biológica, también realizan cálculos complejos para procesar información.

Las aplicaciones de las RNAs incluyen problemas de clasificación y reconocimiento de patrones, imágenes, señales, etc. Asimismo se han utilizado para encontrar patrones de fraude económico, hacer predicciones en el mercado financiero, hacer predicciones de tiempo atmosférico, etc. También se pueden utilizar cuando no existen modelos matemáticos precisos o algoritmos con complejidad razonable, por ejemplo la red de Kohonen [8] ha sido aplicada con un éxito más que razonable al clásico problema del viajante (un problema para el que no se conoce solución algorítmica de complejidad polinómica). Otro tipo especial de redes neuronales artificiales se ha aplicado en conjunción con

los algoritmos genéticos (AG) para crear controladores para robots. La disciplina que trata la evolución de redes neuronales mediante algoritmos genéticos se denomina Robótica Evolutiva.

La computación convencional se caracteriza por el desarrollo de una formación matemática del problema, el desarrollo de un algoritmo para implementar una solución, la codificación del mismo para un problema específico y por último la ejecución de dicho código. Como se ha observado, este tipo de procesamiento es muy exitoso para resolver modelos matemáticos complejos y de simulación, para realizar tareas repetitivas, rápidas y bien definidas. Por otro lado, la computación biológica (derivada del procesamiento en sistemas biológicos) se caracteriza por ser masivamente paralela, adaptativa, lenta, altamente interconectada y tolerante al ruido en el medio ambiente y en sus componentes. De acuerdo con la prensa, publicaciones y conferencias, las redes neuronales (como parte de la computación biológica) han tenido aplicaciones en el área de procesamiento de imágenes y visión computacional [9], específicamente en el análisis de reconocimiento de patrones. Es por ello que en esta tesis el algoritmo a probar estará basado en RNAs.

## 1.1. Objetivos

### 1.1.1. Objetivo general

Implementar una red neuronal para analizar el espectro, obtenido con un dispositivo móvil, de una muestra líquida.

### 1.1.2. Objetivos específicos

1. Comprender los fundamentos teóricos y la técnica experimental para crear una rejilla holográfica

2. Comprender el fenómeno de difracción por una rejilla holográfica
3. Entender los principios de espectroscopía óptica y los componentes de un espectrofotómetro UV-Vis
4. Entender el concepto de las redes neuronales
5. Implementar una red neuronal en espectroscopía
6. Construir un espectrofotómetro con un dispositivo móvil

## 1.2. Descripción de capítulos

La distribución de los capítulos en esta tesis es como sigue. En el capítulo 2 se plantea la teoría de la interferencia desarrollando las ecuaciones necesarias para explicar el concepto físicamente. En el capítulo 3 se inicia con el concepto de difracción seguido de los fundamentos de una rejilla de difracción. En el capítulo 4 se habla de la espectroscopía y sus consideraciones para analizar los datos así como también se analiza la composición y funcionamiento de un espectrofotómetro así como de sus componentes. En el capítulo 5 se dan los fundamentos de una red neuronal, así como su procedimiento que se debe seguir para la implementación del código. En el capítulo 6 se explica a detalle la construcción de los componentes del espectrofotómetro con un dispositivo móvil así como la implementación de la red neuronal. Por último, en el capítulo 7 se dan las conclusiones principales de la tesis.



# Capítulo 2

## Interferencia

### 2.1. Ecuaciones de Maxwell

En 1865, James Clerk Maxwell, un científico que nació en 1831 en el Reino Unido y murió en 1879, publicó un conjunto de 4 ecuaciones, ahora conocidas como ecuaciones de Maxwell, que describen por completo la radiación electromagnética. Ésta consiste en una combinación de campos eléctrico y magnético oscilantes que se propagan a través del espacio vacío, transportando energía de un lugar a otro. Tanto el campo eléctrico como el campo magnético oscilan en direcciones perpendiculares entre sí y a la dirección en que la radiación avanza.

La radiación electromagnética puede manifestarse de diversas maneras tales como ondas de radio, microondas, calor radiado, luz visible, rayos X o rayos gamma. La luz visible está formada por ondas cuya distancia entre dos máximos consecutivos, conocida como longitud de onda, tiene un valor entre los 400 y 700 nanómetros.

Todos los fenómenos electromagnéticos, incluyendo la propagación de la luz, pueden ser descritos totalmente en términos de

las ecuaciones de Maxwell

$$\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0}, \quad (2.1)$$

$$\nabla \cdot \vec{B} = 0, \quad (2.2)$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \quad (2.3)$$

$$\nabla \times \vec{B} = \mu_0 \left( \vec{J} + \epsilon_0 \frac{\partial \vec{E}}{\partial t} \right), \quad (2.4)$$

donde  $\mu_0$  y  $\epsilon_0$  son la permeabilidad y permitividad del vacío;  $\rho$  y  $\vec{J}$  son la densidad de carga y corriente respectivamente.

Se puede demostrar que para el vacío y sin fuentes las ecuaciones de Maxwell se pueden transformar en dos expresiones vectoriales extremadamente concisas [10] y se pueden escribir de la siguiente manera:

$$\nabla^2 \vec{E} = \epsilon_0 \mu_0 \frac{\partial^2 \vec{E}}{\partial t^2}, \quad (2.5)$$

$$\nabla^2 \vec{B} = \epsilon_0 \mu_0 \frac{\partial^2 \vec{B}}{\partial t^2}. \quad (2.6)$$

Éstas son las ecuaciones de onda del campo electromagnético en la propagación de la luz. Una comparación entre la Ec. 2.5 y Ec. 2.6 con la ecuación de la onda sugiere que  $\mu_0 \epsilon_0$  es el inverso del cuadrado de la velocidad de propagación de una onda electromagnética. La velocidad en un medio es determinada por su permitividad y permeabilidad. En el vacío la permeabilidad magnética tiene el valor  $\mu_0 = 4\pi * 10^{-7} N s^2 C^{-2}$  y la permitividad eléctrica  $\epsilon_0 = 8,85 * 10^{-12} C^2 N^{-1} m^{-2}$  así, la velocidad para la propagación de las ondas electromagnéticas en el vacío es de  $c = 2,99 * 10^8 \frac{m}{s} = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$  [11].

Las ondas monocromáticas son caracterizadas por tener una sola frecuencia o color. En este caso, se puede resolver la Ec. 2.5

para encontrar el campo  $\vec{E}$ . Los métodos de análisis complejo proporcionan una forma conveniente de implementar este procedimiento, así:

$$\vec{E}(\vec{r}, t) = \vec{E}(\vec{r})e^{-i\omega t}, \quad (2.7)$$

donde  $\omega$  es la frecuencia de la onda y  $\vec{E}(\vec{r})$  es conocido como faser de la onda, recordando que el campo eléctrico es obtenido tomando la parte real de la Ec. 2.7.  $\vec{E}(\vec{r})$  es en general una función con valores complejos, así que el campo eléctrico es proporcional a  $\cos(\omega t + \phi)$  donde  $\phi$  es llamada fase. Usando la Ec. 2.7 en Ec. 2.5 obtenemos:

$$e^{-i\omega t}(\nabla^2 \vec{E} + \omega^2 \epsilon_0 \mu_0 \vec{E}) = 0. \quad (2.8)$$

La ecuación es gobernada por el término correspondiente a la variación espacial del campo eléctrico. El siguiente paso es resolver dicha ecuación, para ello supóngase que en particular  $\vec{E}(\vec{r})$  varía únicamente en dirección  $z$  y es independiente de  $x$  y  $y$ . Entonces se puede reescribir la Ec. 2.8:

$$\frac{d^2 \vec{E}(z)}{dz^2} + \frac{\omega^2}{\mu_0 \epsilon_0} \vec{E} = 0. \quad (2.9)$$

Esta ecuación, que es conocida como la ecuación de Helmholtz, es matemáticamente igual a la del oscilador armónico simple, por lo que se puede afirmar que una solución posible para  $\vec{E}(z)$  es de la forma

$$\vec{E}(z) = \vec{E}_0 e^{\pm ikz}, \quad (2.10)$$

donde  $\vec{E}_0$  es un vector constante y  $k = \omega/c$ . Poniendo la ecuación para  $\vec{E}(\vec{r})$  en Ec. 2.7 se obtiene la solución completa

$$\vec{E}(\vec{r}, t) = \vec{E}_0 e^{i(\pm kz - \omega t)}. \quad (2.11)$$

La Ec. 2.11, es la expresión de una onda viajera que se propaga en la dirección  $z$  con frecuencia  $\omega$  y amplitud  $\vec{E}_0$ . De igual

manera se puede encontrar el campo magnético siguiendo el procedimiento anterior y se puede demostrar que está dado por:

$$\vec{B}(\vec{r}, t) = \vec{B}_0 e^{i(\pm kz - \omega t)}. \quad (2.12)$$

## 2.2. Interferencia

Una de las propiedades más significativas de la onda electromagnética es que transporta energía. Cualquier campo electromagnético existe dentro de alguna región del espacio y es por consiguiente muy natural considerar la energía por unidad de volumen, es decir la densidad de energía  $u$ . Para un campo eléctrico, se obtiene que ésta es

$$u_E = \frac{\epsilon_0}{2} E^2, \quad (2.13)$$

donde  $E^2$  es la magnitud del campo eléctrico al cuadrado. Similarmen- te para el campo magnético es

$$u_B = \frac{1}{2\mu_0} B^2, \quad (2.14)$$

donde  $B^2$  es la magnitud del campo magnético al cuadrado. Como se puede deducir la relación  $E = cB$  específicamente para una onda plana [12], es fácil ver que:

$$u_E = u_B. \quad (2.15)$$

El flujo de energía a través del espacio en la forma de una onda electromagnética es compartido por los campos constitutivos, eléctricos y magnéticos. Ya que

$$u = u_E + u_B, \quad (2.16)$$

de aquí que

$$u = \epsilon_0 E^2. \quad (2.17)$$

El transporte de energía por unidad de tiempo a través de un área unitaria denotado como  $S$  (en el sistema MKS tendrá unidades  $W/m^2$ ), está dado por la siguiente expresión:

$$S = \frac{uc\Delta tA}{\Delta tA} = uc \quad (2.18)$$

y usando la Ec. 2.17 se obtiene

$$S = \frac{1}{\mu_0} EB. \quad (2.19)$$

En el vacío se tiene que la energía fluye en la dirección en que se propaga la luz, el vector  $\vec{S}$  corresponde entonces a

$$\vec{S} = \frac{1}{\mu_0} \vec{E} \times \vec{B}. \quad (2.20)$$

A  $\vec{S}$  se le conoce como el vector de Poynting. Como  $\vec{S}$  es una función del tiempo que varía extremadamente rápido y su valor es impráctico de medir se suelen emplear promedios. El valor promedio del vector de Poynting, simbolizado por  $\langle S \rangle$ , se conoce como irradiancia,  $I$ . Esto es

$$\langle S \rangle = \frac{c^2 \epsilon_0}{2} | \vec{E}_0 \times \vec{B}_0 | \quad (2.21)$$

o

$$I = \langle S \rangle = \frac{c\epsilon_0}{2} E_0^2 = c\epsilon_0 \langle \vec{E}^2 \rangle, \quad (2.22)$$

donde  $\langle \vec{E}^2 \rangle$  es el promedio en el tiempo de la magnitud de la intensidad del campo eléctrico al cuadrado o  $\langle \vec{E} \cdot \vec{E} \rangle$ .

Cuando dos o más ondas de luz se superponen entre sí, el campo resultante, dado por el principio de superposición [10], es la suma

vectorial de los campos asociados con las ondas individuales, esto es

$$\vec{E}(\vec{r}, t) = \sum_i \vec{E}_i(\vec{r}, t). \quad (2.23)$$

Haciendo referencia a la Ec. 2.22 y considerando que nos concierne las irradiancias relativas dentro del mismo medio, despreciemos las constantes y entonces

$$I = \langle E^2 \rangle, \quad (2.24)$$

Considerando sólo dos ondas linealmente polarizadas [12] se obtiene

$$\vec{E}^2 = \vec{E}_1^2 + \vec{E}_2^2 + 2\vec{E}_1 \cdot \vec{E}_2, \quad (2.25)$$

Tomando el promedio en ambos lados, la irradiancia queda

$$I = I_1 + I_2 + I_{12}, \quad (2.26)$$

donde  $I_1 = \langle E_1^2 \rangle$ ,  $I_2 = \langle E_2^2 \rangle$  e  $I_{12} = 2 \langle \vec{E}_1 \cdot \vec{E}_2 \rangle$ . La última expresión se conoce como término de interferencia. Si se sigue con la forma vectorial del campo eléctrico, considerando diferencias de fase  $\epsilon_1$  y  $\epsilon_2$  respectivamente para cada onda, se obtiene [12]

$$I_{12} = E_{01}E_{02} \cos \delta, \quad (2.27)$$

siendo  $\delta = \vec{k}_1 \cdot \vec{r} - \vec{k}_2 \cdot \vec{r} + \epsilon_1 - \epsilon_2$  la diferencia de fase que proviene de combinar una diferencia de longitud de trayectoria y una diferencia de fase inicial. Si se toma  $\vec{E}_1$  paralelo a  $\vec{E}_2$  se simplifican las expresiones para  $I_1$ ,  $I_2$  e  $I_{12}$  obteniendo

$$I_1 = E_{01}^2, I_2 = E_{02}^2, I_{12} = 2\sqrt{I_1 I_2} \cos \delta, \quad (2.28)$$

donde la irradiancia final es

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos \delta. \quad (2.29)$$

En los puntos donde  $\delta$  es múltiplo de  $2\pi$  se obtienen máximos de irradiancia o también se dice que las perturbaciones están en

fase. Se habla de esto como interferencia constructiva total. En el caso en donde  $\delta$  es múltiplo de  $\pi$  se obtiene un mínimo total de irradiancia o las perturbaciones están  $180^\circ$  fuera de fase, ésta es la condición de interferencia destructiva total.

Otro caso especial y muy importante se da cuando las amplitudes de ambas ondas que alcanzan el punto  $P$  en la Fig. 2.1 son iguales y tienen la misma orientación (es decir  $\vec{E}_{01} = \vec{E}_{02}$ ). Puesto que las contribuciones de la irradiancia de ambas fuentes son iguales, sea  $I_1 = I_2 = I_0$ . Lo que permite escribir la Ec. 2.29 como

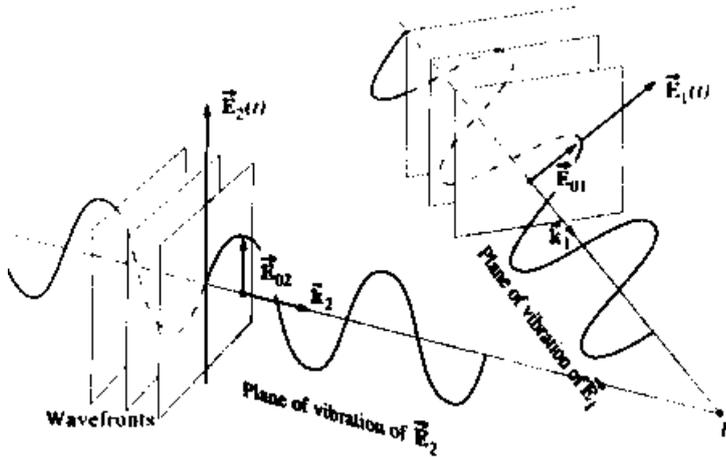


Fig. 2.1: Dos ondas en el espacio superponiéndose en el punto P. Imagen tomada de [12].

$$I = 2I_0(1 + \cos \delta) = 4I_0 \cos^2 \frac{\delta}{2}. \quad (2.30)$$

La Ec. 2.29 también se aplica a ondas esféricas emitidas por  $S_1$  y  $S_2$ . Dichas ondas pueden escribirse como

$$\vec{E}_1(r_1, t) = \vec{E}_{01}(r_1) \exp[i(kr_1 - \omega t + \epsilon_1)], \quad (2.31)$$

$$\vec{E}_2(r_2, t) = \vec{E}_{02}(r_2) \exp[i(kr_2 - \omega t + \epsilon_1)]. \quad (2.32)$$

Las cantidades  $r_1$  y  $r_2$  son los radios de los frentes de onda esféricos que se superponen en P, es decir, ellos especifican las distancias de las fuentes hasta P. De esto que

$$\delta = k(r_1 - r_2) + (\epsilon_1 - \epsilon_2). \quad (2.33)$$

Siguiendo la Ec. 2.30 los máximos de irradiancia ocurren cuando

$$\delta = 2\pi m, \quad (2.34)$$

para  $m$  entero. Del mismo modo, los mínimos aparecen para

$$\delta = \pi m, \quad (2.35)$$

para  $m$  entero y distinto de 0. Ahora se asocia la Ec. 2.33 con 2.34 y si se considera que las ondas están en fase,  $\epsilon_1 - \epsilon_2 = 0$  se obtiene la relación para la irradiancia máxima,

$$r_1 - r_2 = \pi m/k = m\lambda. \quad (2.36)$$

Si la diferencia de fase entre las dos fuentes permanece constante en el tiempo y espacio, tales fuentes se les llama coherentes. Dos haces que se superponen y que vienen de emisores separados interferirán pero el patrón resultante no se sostendrá el tiempo suficiente para ser fácilmente observable. La forma fácil de resolver el problema es hacer que una fuente se utilice para producir dos fuentes secundarias coherentes.

Los patrones más claros existirán cuando las ondas que interfieren tengan amplitudes iguales o casi iguales. Las regiones centrales de las franjas oscuras y claras corresponden entonces a interferencia completamente destructiva o constructiva, respectivamente, dando mayor contraste.

Fresnel y Arago hicieron un estudio extensivo de las condiciones bajo las cuales la interferencia de luz polarizada ocurre y sus conclusiones son llamadas Leyes de Fresnel-Arago y son las siguientes:

- Dos estados  $\mathbf{P}^1$  coherentes ortogonales no pueden interferir. Esto significa que  $I_{12} = 0$  y no resulten franjas.
- Dos estados  $\mathbf{P}$  coherentes y paralelos interfieren en la misma forma que la luz natural.
- Dos estados  $\mathbf{P}$  ortogonales constitutivos de la luz natural no pueden interferir para formar un patrón fácilmente observable aunque se giran para alinearlos. Este último punto es comprensible ya que estos estados  $\mathbf{P}$  son incoherentes.

### 2.3. Interferómetros de dos haces

Para generar interferencia de dos haces existen dispositivos llamados interferómetros que se clasifican en interferómetros de división de frente de onda y división de amplitud. El ejemplo clásico de los interferómetros de división de amplitud es el interferómetro de Young que se describe a continuación.

Considérese una onda plana monocromática hipotética que ilumina una rendija larga y estrecha. De esa rendija primaria, la luz se difractará con todos los ángulos hacia delante y emergerá una onda cilíndrica. Supóngase que esta onda, a su vez, incida en dos rendijas  $S_1$  y  $S_2$  muy juntas, estrechas y paralelas tal como se muestra en la Fig. 2.2a. Cuando existe simetría, los segmentos del frente de onda primario que llegan a las dos rendijas estarán exactamente en fase y las rendijas constituirán dos fuentes secundarias coherentes. Es de suponer que donde quiera que las dos ondas procedentes de  $S_1$  y  $S_2$  se superpongan, se producirá interferencia.

La Fig. 2.2b corresponden a la disposición clásica del *experimento de Young* si bien existen otras variantes. La diferencia de

---

<sup>1</sup>El estado  $\mathbf{P}$  se refiere a la polarización que tienen las ondas incidentes, para una explicación más amplia véase [12].

camino óptico entre los rayos a lo largo de  $\overline{S_1P}$  y  $\overline{S_2P}$  (véase Fig. 2.2c.) puede calcularse bajando una perpendicular desde  $S_2$  hasta  $\overline{S_1P}$ . Esta diferencia de camino óptico viene dada por

$$\overline{S_1B} = \overline{S_1P} - \overline{S_2P} \quad (2.37)$$

o

$$\overline{S_1B} = r_1 - r_2. \quad (2.38)$$

Continuando con esta aproximación,  $r_1 - r_2 = a \sin \theta$  y por lo tanto

$$r_1 - r_2 \approx a\theta, \quad (2.39)$$

puesto que  $\theta \approx \sin \theta$  ya que  $\theta \ll 1$ . Por la Ec. 2.36 la interferencia constructiva ocurrirá cuando  $r_1 - r_2 = m\lambda$  entonces obtenemos

$$y_m \approx \frac{s}{a}m\lambda. \quad (2.40)$$

Esto proporciona la posición de la  $m$ -ésima franja brillante en la pantalla si se considera el máximo en 0 como la franja de orden cero. La posición angular de las franjas se obtiene substituyendo la última expresión en  $\theta \approx \frac{y}{s}$ ; así

$$\theta_m = \frac{m\lambda}{a}. \quad (2.41)$$

En el espacio entre las franjas en la pantalla puede obtenerse fácilmente de la Ec. 2.40. La diferencia en las posiciones de dos máximos consecutivos es

$$\begin{aligned} \Delta y = y_{m+1} - y_m &\approx \frac{s}{a}(m+1)\lambda - \frac{s}{a}m\lambda \\ \Delta y &\approx \frac{s}{a}\lambda. \end{aligned} \quad (2.42)$$

El interferómetro típico empleado como ejemplo de los interferómetros de división de amplitud es el interferómetro de Michelson, sin embargo en esta tesis consideramos el interferómetro de dos haces que se muestra en la Fig. 2.3.

En la Fig. se observa un haz de luz proveniente de un láser que incide sobre un semiespejo o divisor de haz que transmite el 50 % de luz hacia el espejo e1 y refleja el restante 50 % hacia el espejo e2. Los espejos e1 y e2 se colocan a la misma distancia del divisor de haz, separados una distancia  $a$  entre ellos y se orientan de tal forma que reflejen la luz hacia un punto  $p$  situado sobre una pantalla alejada una distancia  $s$  del punto medio que une los espejos. Como se observa en la Fig. 2.3, la geometría de este interferómetro es similar a la geometría del interferómetro de Young (Fig. 2.2c) por lo que la matemática para encontrar el patrón de interferencia en este instrumento es similar a la descrita anteriormente.

Como resultado de la superposición de los haces en el arreglo experimental descrito en la Fig. 2.3 se obtendrá un patrón de interferencia en forma de franjas claras y oscuras alternadas. Si ahora sobre la pantalla de observación se coloca un material fotosensible, como una película holográfica, el patrón de interferencia se grabará sobre la superficie del material. La imagen que se registrará sobre la película corresponderá a una rejilla de difracción cuyas características y elaboración se discutirán en el capítulo siguiente.

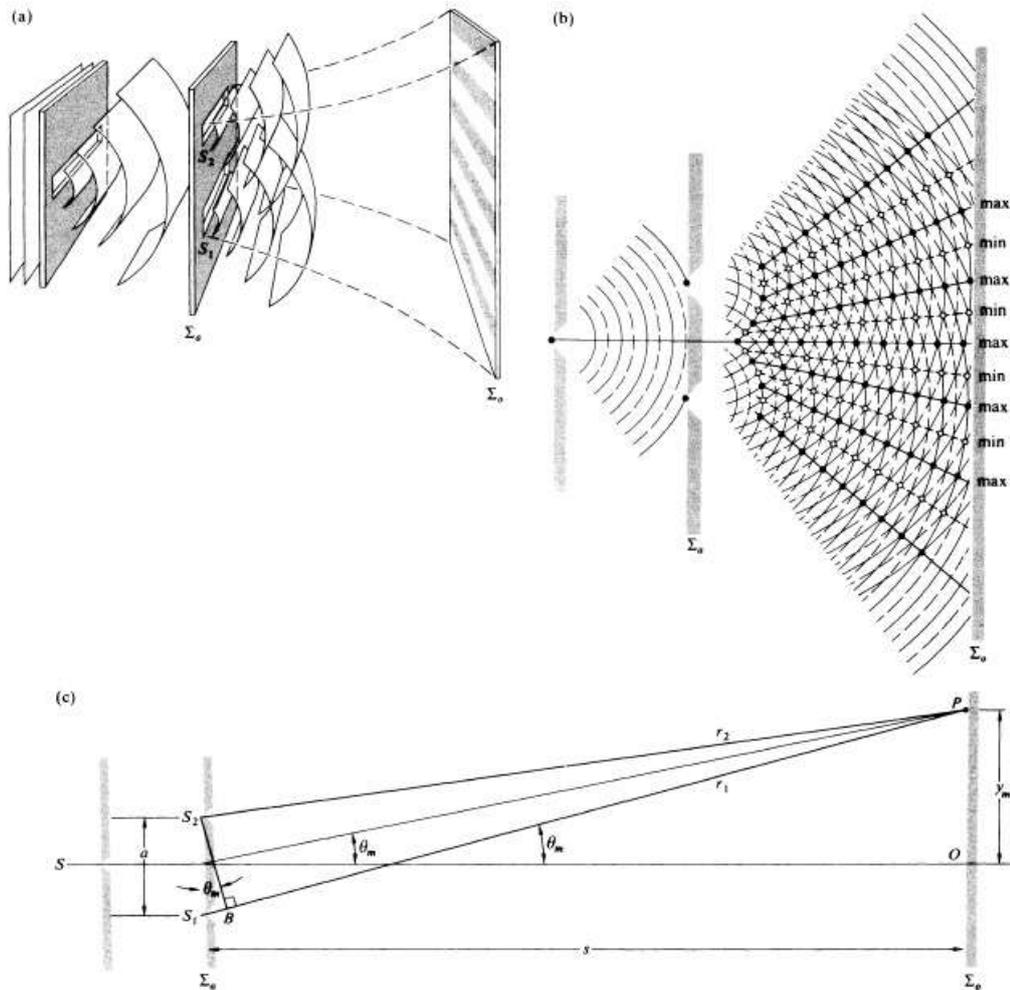


Fig. 2.2: (a) Ondas cilíndricas superponiéndose en la región posterior a la pantalla donde están las aperturas. (b) Descripción gráfica de la superposición de ondas en la disposición clásica del experimento de Young. (c) Geometría del experimento de Young. Imagen tomada de [12].

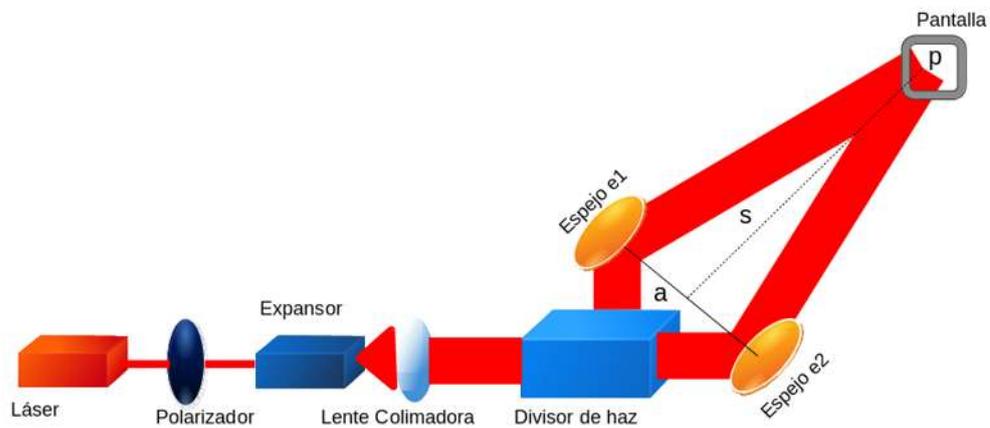


Fig. 2.3: Esquema del interferómetro de dos haces utilizado en esta tesis. La amplitud es dividida y reflejada por los espejos  $e1$  y  $e2$  hacia la pantalla donde generan un patrón de interferencia.



## Capítulo 3

# Difracción y rejillas holográficas

Un cuerpo opaco colocado a medio camino entre una pantalla y una fuente puntual proyecta una sombra complicada hecha de regiones claras y oscuras, muy diferentes de las que podría esperarse de los principios de la óptica geométrica. El trabajo de Francesco Grimaldi en el siglo XVII fue el primer estudio detallado que se publicó sobre esta desviación de la luz de su propagación rectilínea, a la que denominó *diffraçtio*. El efecto es una característica general de los fenómenos ondulatorios que ocurren donde quiera que una parte del frente de onda esté obstruida de alguna manera. Si al encontrar un obstáculo transparente u opaco se altera la amplitud o la fase de una región del frente de onda, esto producirá difracción. Los varios segmentos del frente de onda que se propagan más allá del obstáculo interfieren, produciendo aquella distribución de densidad de energía particular denominada patrón de difracción. No hay distinción física significativa entre interferencia y difracción. Sin embargo, se ha vuelto algo común, aunque no siempre apropiado, hablar de interferencia cuando se analiza la superposición de solamente unas pocas ondas y de difracción cuando se trata de un gran número de ondas.

### 3.1. Principio de Huygens-Fresnel

El principio de Huygens establece que cada punto en un frente de onda en propagación sirve como fuente de trenes de ondas esféricas secundarias de tal modo que, al cabo de cierto tiempo, el frente de onda será la envolvente de estos trenes de ondas. Por otro lado, si la onda que se propaga tiene una frecuencia  $\nu$  y se transmite por el medio a una velocidad  $v_t$ , entonces los trenes de ondas secundarios tendrán la misma frecuencia y velocidad.

Como cada punto en un frente de onda puede visualizarse como una fuente de trenes de ondas esféricas secundarias, se podría supuestamente determinar el progreso del frente de onda de cualquier porción de ella a través del espacio. Sin embargo, no toma en consideración la mayoría de los trenes de onda secundarias, reteniendo solo la parte común a la envolvente. Como resultado de esta imperfección el principio de Huygens no puede explicar, por sí mismo, los detalles del proceso de difracción.

El principio de Huygens-Fresnel establece que cada punto sin obstrucción de un frente de onda, en un instante de tiempo determinado, sirve como fuente de trenes de onda secundarias esféricas (de la misma frecuencia que la onda primaria). La amplitud del campo óptico en cualquier punto más allá es la superposición de todos estos trenes de onda (considerando sus amplitudes y fases relativas).

Considere la Fig. 3.1, si cada punto despejado de la onda plana incidente actúa como fuente secundaria coherente, la máxima diferencia en las longitudes de camino óptico entre ellas será  $\Lambda_{max} = \| \overline{AP} - \overline{BP} \|$ , correspondiendo a una fuente puntual en cada borde de la abertura. Pero  $\Lambda_{max} \leq \overline{AB}$ , siendo igual cuando  $P$  se encuentra sobre la pantalla. Cuando  $\lambda > \overline{AB}$  se deduce que  $\lambda > \Lambda_{max}$  y puesto que las ondas estaban inicial-

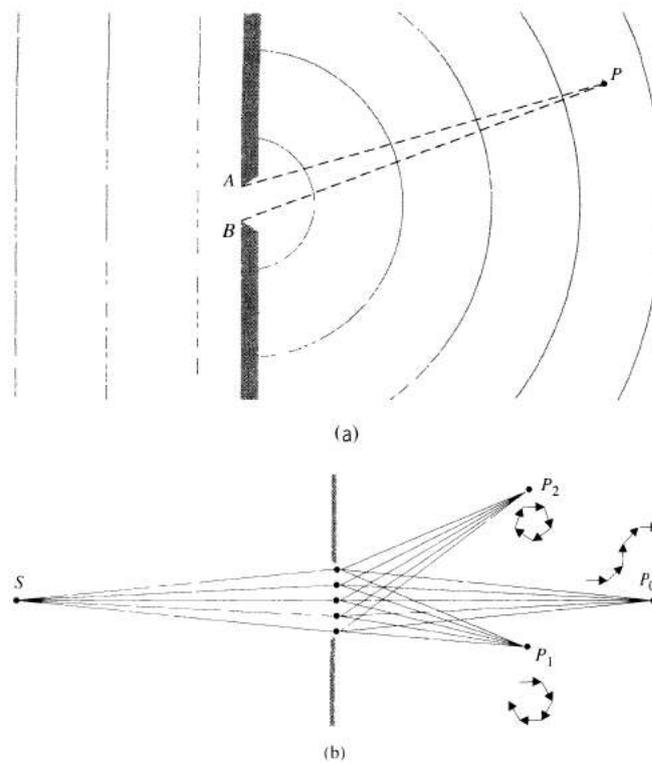


Fig. 3.1: Difracción en una abertura pequeña. (a) Imagen de onda clásica. (b) Vista realizada con EDC y amplitudes de probabilidad. Imagen tomada de [12].

mente en fase, interfieren constructivamente (en grado variable) donde quiera que esté  $P$ . Por lo tanto, si la longitud de onda es grande comparada con la abertura, las ondas se extenderán según ángulos grandes en la región más allá de la obstrucción.

Desde un punto de vista clásico, la razón por la que la luz va más allá de la pantalla se debe a la *interferencia* de la multitud de ondas secundarias que se emiten de la abertura; es decir, todas ellas se combinan en cada punto de la región, en algunos reforzándose, en otros anulándose, según la diferencia de camino óptico.

### 3.2. Difracción de Fraunhofer y Fresnel

Supongase una pantalla opaca  $\Sigma$  que tiene una pequeña abertura iluminada por una onda plana de una fuente puntual,  $S$ , muy lejana. El plano de observación  $\sigma$  es una pantalla paralela y muy cercana a  $\Sigma$ . En estas condiciones una imagen de la abertura se proyecta sobre la pantalla, que es claramente reconocible a pesar de una ligera franja alrededor de su periferia (Fig. 3.2). Si el plano de observación se mueve más lejos de  $\Sigma$ , la imagen de la abertura, aunque todavía fácilmente reconocible, se hace cada vez más estructurada como las franjas se vuelven más prominentes. Este fenómeno se conoce como difracción de Fresnel o de campo cercano. Si el plano de la observación se aleja aún más, se produce un cambio continuo en las franjas. A una distancia muy grande de  $\Sigma$  el patrón proyectado cambió considerablemente, teniendo poco o ningún parecido con la abertura real. A partir de entonces cambia sólo el tamaño del patrón y no su forma. Esta es la difracción de Fraunhofer o de campo lejano. Si en ese punto se pudiera reducir suficientemente la longitud de onda de la radiación entrante, el patrón volverá al caso de Fresnel. Si la distancia se reduce aún más, de manera que se aproxime a cero, las franjas desaparecerían, y la imagen tomaría la forma limitadora de la abertura, como predice la óptica geométrica. Volviendo a la configuración original, si la fuente puntual se mueve ahora hacia  $\Sigma$ , algunas ondas esféricas podrán incidir en la abertura, y no existiría un patrón de Fresnel, incluso en un plano distante de la observación.

Si las ondas entrantes y salientes se aproximan a ser planas durante la extensión de las aberturas de difracción (u obstáculos), se obtiene la difracción de Fraunhofer. Otra forma de apreciar esto es darse cuenta de que la fase de cada contribución a  $P$ , debido a diferencias en el camino recorrido, es crucial para la

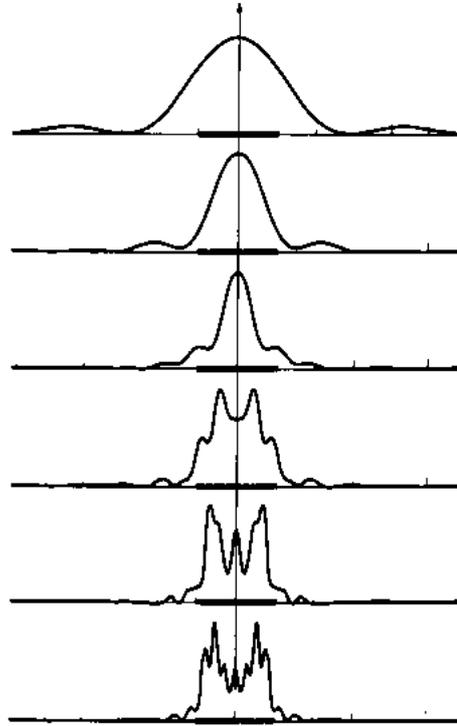


Fig. 3.2: Una sucesión de distribuciones de difracción a distancias crecientes de una rendija única. Imagen tomada de [12].

determinación del campo resultante. Por otra parte, si los frentes de onda que inciden en, y salen de, la abertura son planas, entonces estas diferencias de camino serán descritas por una función lineal de las dos variables de apertura. Esta linealidad en las variables de apertura es el criterio matemático definitivo de la difracción de Fraunhofer. Por otro lado, cuando  $S$  o  $P$ , o ambos están demasiado cerca de  $\Sigma$  como para considerar la curvatura de los frentes de onda entrantes y salientes insignificantes, la difracción de Fresnel prevalece.

Por regla empírica, la difracción de Fraunhofer se producirá en una abertura (u obstáculo) con largo máximo  $a$  cuando

$$R > a^2/\lambda, \quad (3.1)$$

donde  $R$  es la distancia más pequeña de las dos que van de  $S$

hasta  $\Sigma$  y de  $\Sigma$  hasta  $P$ . Cuando  $R \rightarrow \infty$  la dimensión finita de la abertura es de poco interés. Así mismo, un aumento de  $\lambda$  desplaza claramente el fenómeno hacia el extremo de Fraunhofer.

Una realización práctica de la condición de Fraunhofer, cuando ambos  $S$  y  $P$  están efectivamente en el infinito, se logra recurriendo a una disposición como la Fig. 3.3. La fuente puntual  $S$  se ubica en  $F_1$ , el foco principal de lente  $L_1$  y el plano de observación es el segundo plano focal de  $L_2$ . Según la terminología de la óptica geométrica, el plano de la fuente y  $\sigma$  son planos conjugados.

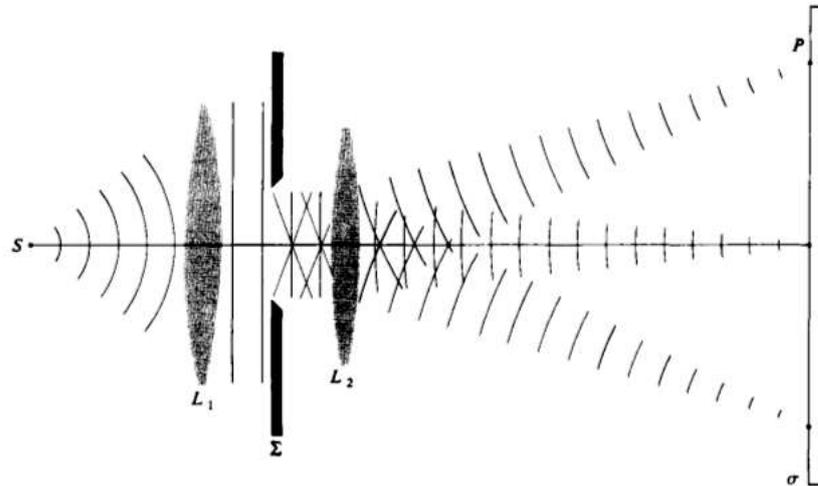


Fig. 3.3: Difracción de Fraunhofer. Imagen tomada de [12].

### 3.3. Rejillas de difracción

Una rejilla de difracción es la colección de elementos de reflexión (o de transmisión) separados por una distancia comparable con la longitud de onda de la luz en estudio. Pueden ser pensados como una colección de elementos de difracción, tales como un patrón de hendiduras (también llamadas aperturas o

simplemente líneas) transparentes en una pantalla opaca, o una colección de ranuras que reflejan en un sustrato. Una rejilla de reflexión consiste en una rejilla superpuesta sobre una superficie reflectante, mientras que una rejilla de transmisión consiste en una rejilla superpuesta sobre una superficie transparente. En cualquiera de los casos anteriores, las rejillas de difracción tienen como características fundamentales la modulación del índice de refracción, modulación del relieve o modulación de su función de transmitancia lo que ocasiona una modulación en la amplitud de la onda incidente. Tras la difracción, una onda electromagnética incidente en una rejilla tendrá su amplitud de campo eléctrico, o fase, o ambas, modificadas de una manera predecible, debido a la variación periódica en la región cerca de la superficie de la rejilla.

### 3.3.1. Ecuación de la rejilla

Cuando un haz de luz monocromático incide en la superficie de la rejilla, éste es difractado en direcciones discretas. La luz difractada por cada hendidura se combina a partir de un frente de onda difractado. La utilidad de una rejilla depende del hecho de que existen un conjunto único de ángulos discretos a lo largo de los cuales, para una separación  $d$  dada entre las ranuras, la luz difractada de cada hendidura está en fase con la luz difractada desde cualquier otra hendidura, de modo que se combinan constructivamente.

La difracción por una rejilla puede ser vista por una geometría como en la Fig. 3.4a, la cual muestra un rayo de luz con longitud de onda  $\lambda$  incidiendo en un ángulo  $\alpha$  y difractado por una rejilla (por un espacio  $d$  entre hendiduras, también llamado paso) a lo largo de  $\beta_m$  ángulos. Estos ángulos son medidos respecto a la normal de la rejilla, ésta es indicada por la línea perpendicular

a la superficie de la rejilla. La convención de signos depende de si la luz es difractada sobre el mismo lado o el lado opuesto de la superficie donde incide la luz. En la Fig. 3.4a se muestra una rejilla de reflexión, donde  $\alpha, \beta_1 > 0$  y  $\beta_0, \beta_{-1} < 0$ . En la Fig. 3.4b se muestra una rejilla de transmisión.

Por convención, los ángulos incidentes y difractados son medidos sobre la normal de la rejilla al haz. En ambos diagramas, la convención para los ángulos se sigue del símbolo más o menos situados a ambos lados de la normal de la rejilla en la Fig. 3.4. Para ambas rejillas de difracción o de transmisión, los signos de los ángulos difieren si ellos son medidos sobre lados opuestos de la normal de la rejilla.

La relación entre los ángulos es dada por la ecuación de la rejilla

$$m\lambda = d(\sin \alpha + \sin \beta), \quad (3.2)$$

la cual gobierna los ángulos de difracción para una rejilla de separación entre hendiduras  $d$ . Aquí  $m$  es el orden de difracción, el cual es un entero. Para una onda de longitud de onda  $\lambda$ , todos los valores de  $m$  obedecen la relación  $\|m\lambda/d\| < 2$  que corresponde a los órdenes de difracción físicamente posibles.

En muchas ocasiones se hace mención a  $G = 1/d$  como la frecuencia de hendiduras o densidad de hendiduras, más comúnmente *líneas por milímetro*. Por lo tanto, se puede reescribir la Ec.3.2 como

$$Gm\lambda = \sin \alpha + \sin \beta. \quad (3.3)$$

Un caso especial pero común es en el cual la luz es difractada en la dirección en la cual proviene (i.e.,  $\alpha = \beta$ ), este caso es llamado *Littrow configuration*, por lo cual la Ec. 3.2 se escribe como

$$m\lambda = 2d \sin \alpha. \quad (3.4)$$

### 3.3.2. Órdenes de difracción

Generalmente varios enteros  $m$  satisfacen la ecuación de la rejilla y se nombra a cada entero como un orden de difracción.

Para una ranura particular de espaciamiento  $d$ , de la longitud de onda  $\lambda$  y de ángulo incidente  $\alpha$ , la ecuación de la rejilla (Ec. 3.2) es generalmente satisfecha por más de un ángulo de difracción  $\beta$ . De hecho, sujeto a las restricciones mencionadas a continuación, habrá varios ángulos discretos en el que se cumple la condición de interferencia constructiva. El significado físico de esto es que el refuerzo constructivo de las longitudes de onda difractadas por las ranuras sucesivas sólo exige que cada rayo sea retardado (o avanzado) en fase con cada otra; por lo tanto, esta diferencia de fase debe corresponder a una distancia real (diferencia de camino) que es igual a un múltiplo entero de la longitud de onda. Esto sucede, por ejemplo, cuando la diferencia de camino es una longitud de onda, en cuyo caso se habla del primer orden de difracción positivo ( $m = 1$ ) o el primer orden de difracción negativo ( $m = -1$ ), dependiendo de si los rayos son avanzados o retardados a medida que pasamos de una ranura a otra ranura. Del mismo modo, el segundo orden ( $m = 2$ ) y segundo orden negativo ( $m = -2$ ) son aquellos para los que la diferencia de caminos entre los rayos difractados de ranuras adyacentes es igual a dos longitudes de onda.

La Ec. 3.2 revela que únicamente los órdenes espectrales para los cuales  $|m/d| < 2$  pueden existir; de lo contrario,  $|\sin \alpha + \sin \beta| > 2$ , lo cual no tiene físicamente sentido. Esta restricción evita que la luz de longitud de onda  $\lambda$  sea difractada en más de un número finito de órdenes. La reflexión especular ( $m = 0$ ) es siempre posible; es decir, siempre existe el orden cero (que simplemente requiere  $\beta = -\alpha$ ). En la mayoría de los casos, la ecuación de la rejilla permite que la luz de longitud de onda  $\lambda$  sea difracta-

da en ambos órdenes negativos y positivos. Explícitamente, los espectros de todos los órdenes para los que  $m$  que existe

$$-2d < m\lambda < 2d. \quad (3.5)$$

Si la luz incidente en la rejilla es luz blanca<sup>1</sup>, cada longitud de onda se abrirá a un distinto ángulo según Ec. 3.2, en consecuencia para cualquier orden  $m$  se verán los colores correspondientes al espectro de la luz visible que se discutirá en el siguiente capítulo.

---

<sup>1</sup>La luz blanca es una superposición de luces de diferentes colores las cuales presentan una longitud de onda y una frecuencia específicas.

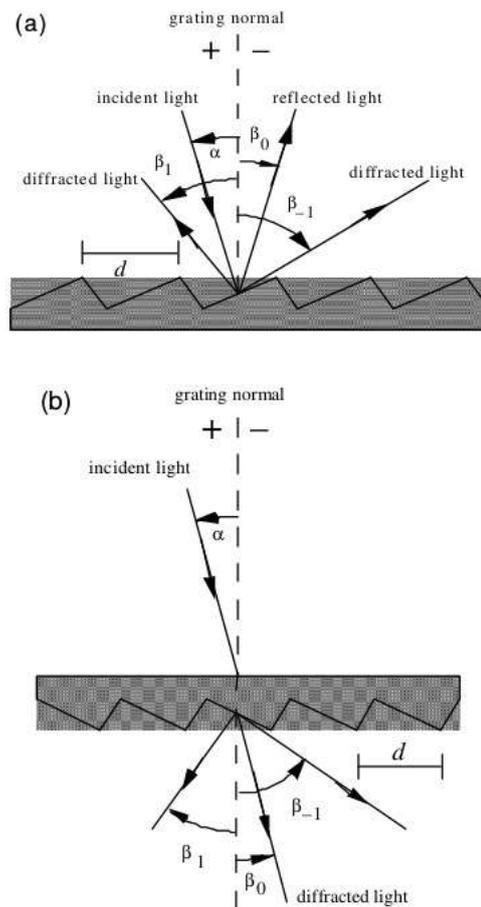


Fig. 3.4: (a) Rejilla de reflexión: la luz incidente y difractada se encuentran en el mismo lado de la rejilla. (b) Rejilla de transmisión: la luz difractada se encuentra en el lado opuesto de la rejilla respecto a la luz incidente. Imagen tomada de [6].



# Capítulo 4

## Espectroscopía

La espectroscopía es el estudio del espectro electromagnético, la composición de la luz en longitudes de onda, debido a las interacciones atómicas y moleculares. Por muchos años, la espectroscopía ha sido de importancia en el estudio de la física, y es ahora de igual importancia en astronomía, biología, química y otras áreas de investigación analítica.

La espectroscopía describe la interacción entre la radiación electromagnética y la materia, tal interacción se lleva a cabo con los cuantos de la radiación electromagnética y estos siempre están caracterizados por una longitud de onda  $\lambda$ , una frecuencia  $\nu$  o una energía  $E$ . La relación entre ellas viene dada por la ecuación de Plank:

$$E = h\nu = hc/\lambda. \quad (4.1)$$

En la Fig. 4.1 se da un esquema de los distintos valores de frecuencia y longitud de onda y la región del espectro que les corresponde.

La radiación electromagnética ocasiona distintos efectos en la materia, tal como se muestra en la tabla 4.1.

Cuando la radiación incide sobre una sustancia, sólo un átomo

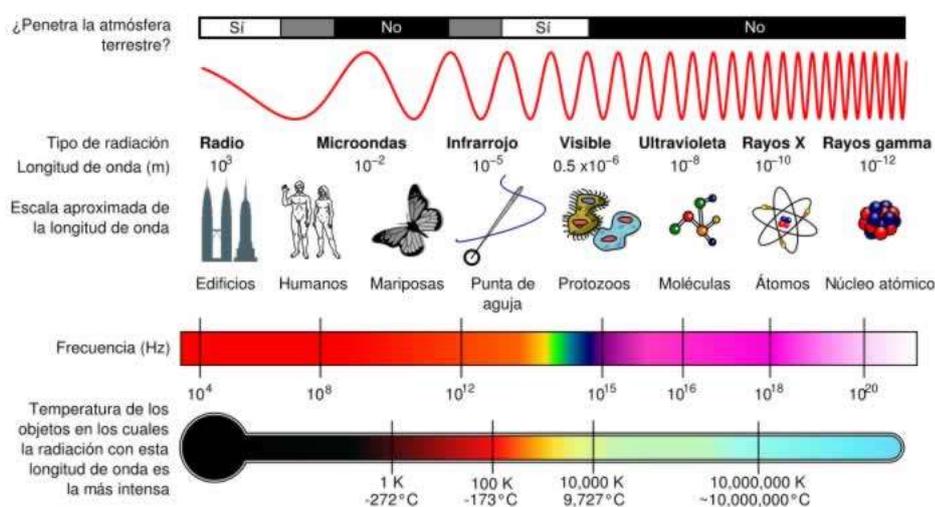


Fig. 4.1: Relación de los valores de frecuencia y longitud de onda de las distintas regiones del espectro de la luz. Imagen tomada de [2].

o un conjunto de átomos son capaces de absorber la radiación; estos grupos se denominan *cromóforos* y serán distintos dentro de una misma molécula para cada técnica espectroscópica.

Los efectos de la radiación sobre la materia pueden usarse para obtener información sobre la estructura de la misma, así surgen distintas técnicas espectroscópicas tal como se muestra en la tabla 4.2.

Tabla 4.1: Efecto de la radiación electromagnética en la materia.

Radiación	Efecto
Rayos X y cósmicos	Ionizaciones de las moléculas
UV - Vis	Transiciones electrónicas entre los orbitales atómicos y moleculares
Infrarrojo	Deformación en los enlaces químicos
Microondas	Rotaciones de los enlaces químicos
Radiofrecuencia	Transiciones de espín electrónico o nuclear en los átomos de la molécula

Tabla 4.2: Técnicas espectroscópicas e información que se puede obtener de ellas.

Técnica espectroscópica	Información obtenida
Rayos X	Estructura total de la molécula incluida la estereoquímica
UV - Vis	Existencia de cromóforos y/o conjugación en la molécula
Infrarrojo	Grupos funcionales a partir de las absorciones observadas
Espectroscopía de masas	Fórmula molecular y subestructuras a partir de los iones observados
Resonancia magnética nuclear	Grupos funcionales, subestructuras, conectividades, estereoquímica, etc.

## 4.1. Espectrometría Ultravioleta Visible (UV-Vis)

La espectroscopía UV-Vis utiliza la radiación del espectro electromagnético, cuya longitud de onda está entre los 100 y los 800 nm y su efecto sobre la materia es producir transiciones electrónicas entre los orbitales atómicos y/o moleculares de la sustancia. En algunos casos, los efectos y la detección pueden llegar al IR cercano (800-900nm). En la espectroscopía UV-Vis una especie química (en general una molécula, aunque puede tratarse de una especie monoatómica, un ión o un complejo) absorbe UV-Vis, y la energía adquirida por el sistema causa la transición de un electrón de un estado basal o fundamental (EF) a uno excitado (EE). La energía de transición está relacionada con la longitud de onda de radiación a través de la ecuación de Planck (Ec. 4.1).

Un gráfico o representación de la respuesta del sistema en función de la longitud de onda o frecuencia se denomina espectro. En general, en los espectros UV-Vis, se observa una señal de-

bida a cada transición electrónica del EF al EE. Los átomos dan líneas agudas, mientras que las moléculas poliatómicas dan señales en forma de bandas puesto que la absorción de luz involucra también energía suficiente para causar cambios en energía vibracional y rotacional de cada uno de sus estados electrónicos en el EE. Para una sustancia determinada, la longitud de onda a la cual se produce el máximo de absorbancia en el espectro se denotará como  $\lambda_{max}$ .

La señal espectral permite, por un lado, identificar algunos grupos funcionales presentes en las moléculas y, por el otro lado, estimar la concentración de una sustancia. La espectrometría es la técnica espectroscópica usada para evaluar la concentración de una especie y utiliza un instrumento llamado espectrómetro. En el caso de la espectrometría que utiliza fotones (UV-Vis, IR), se suele hablar de espectrofotometría. Para la medición de la intensidad de absorción se usan espectrofotómetros en los cuales se puede medir la absorbancia o la transmitancia.

## 4.2. Ley de Beer-Lambert

La ley de Lambert hace mención en que la proporción de radiación que una sustancia absorbe, es independiente de la radiación incidente. Esto es que a cada espesor  $dx$  de una sustancia absorbente le corresponde una fracción igual a  $\frac{dI}{I}$  de la intensidad incidente sobre ella, es decir:

$$\frac{-dI}{I} = b \cdot dx, \quad (4.2)$$

siendo  $b$  una constante y tomando a  $l$  como la longitud de la muestra se sigue el desarrollo de la ecuación diferencial de primer orden (Ec. 4.2). Tomando en cuenta la condición inicial  $I(l =$

$I = I_0$  se obtiene:

$$I = I_0 \exp^{-bl}. \quad (4.3)$$

Es más frecuente utilizar logaritmos, por lo que la Ec. 4.3 se puede escribir de la siguiente manera

$$-\log_{10} \left( \frac{I}{I_0} \right) = \frac{bl}{2,303} = A, \quad (4.4)$$

donde  $A$  es llamada absorbancia.

Beer estudió como influye la concentración de una sustancia sobre la absorbancia y encontró la misma relación lineal entre la absorbancia y el espesor de la muestra. Para una sustancia de concentración  $c$ , la absorbancia está dada por[2]

$$A = c \cdot \epsilon \cdot l, \quad (4.5)$$

donde  $\epsilon$  es una constante conocida como coeficiente de absorción molar. La ley de Beer-Lambert, también conocida como Lambert-Beer, combina las Ecs. 4.5 y 4.4 en una sola expresión:

$$A = \log_{10} \left( \frac{I_0}{I} \right) = c \cdot \epsilon \cdot l. \quad (4.6)$$

Esta expresión determina la absorbancia, es decir la relación entre la intensidad de luz transmitida y la intensidad incidente, con parámetros de la muestra.

### 4.3. El espectrofotómetro UV-Vis

Para las determinaciones analíticas es solamente necesario disponer de un espectrofotómetro UV-Vis y con este equipo medir

el cambio de la absorbancia de un compuesto a diferentes concentraciones de un analito<sup>1</sup> por un método estandarizado para la especie en cuestión. Generalmente, se realiza una curva de calibración de la cual se puede obtener el coeficiente de absorción molar.

La espectrometría UV-Vis se emplea generalmente en la determinación cuantitativa de la concentración en solución de especies químicas como iones metálicos de transición y compuestos orgánicos muy conjugados. Muchas de las determinaciones incluyen un paso de reacción entre la especie y un compuesto que origine un derivado coloreado o absorbente en el UV-Vis.

Por lo general, se usa el agua como disolvente para compuestos inorgánicos y etanol para compuestos orgánicos porque este alcohol absorbe muy débilmente a la mayoría de las longitudes de onda.

El instrumento usado en la espectrofotometría ultravioleta-visible se denomina espectrofotómetro UV-Vis y permite comparar la radiación absorbida o transmitida por una solución que contiene una cantidad desconocida de soluto con una que contiene una cantidad de la misma sustancia. Se mide la transmitancia de la muestra que se expresa habitualmente como porcentaje ( $\%T$ ) o bien la absorbancia ( $A$ ).

### 4.3.1. Partes básicas de un espectrofotómetro

El espectrofotómetro se compone generalmente de una fuente de luz, por lo general una lámpara incandescente (de tungsteno) para las longitudes de onda en el rango visible, o una lámpara de arco de deuterio en el ultravioleta, sin embargo, las lámparas

---

<sup>1</sup>En química analítica, un analito es un componente (elemento, compuesto o ión) de interés analítico de una muestra. Es una especie química cuya presencia o contenido se desea conocer.

de mercurio-xenón también son empleadas, un soporte para la muestra, una rejilla de difracción o monocromador para separar las diferentes longitudes de onda de la luz y un detector. El detector suele ser un fotodiodo o un CCD (*Charge-coupled device*). Los fotodiodos se usan con monocromadores que filtran la luz, de modo que una sola longitud de onda alcanza el detector. Las rejillas de difracción se utilizan en conjunto con CCDs, que recogen radiación electromagnética de diferente longitudes de onda en pixeles.

Un espectrofotómetro puede ser de haz simple o de doble haz. En un instrumento de un sólo haz, toda la luz pasa a través de la celda de la muestra. La intensidad incidente  $I_a$  se mide análogamente en ausencia de la muestra. En un instrumento de doble haz, la luz se divide en dos haces antes de llegar a la muestra. Algunos instrumentos de doble haz tienen dos detectores y se puede medir simultáneamente tanto el haz de referencia como el de la muestra. En otros instrumentos, los dos haces pasan a través de un bloqueador que impide el paso de uno de los haces. El detector alterna entre la medida del haz de muestra y la del haz de referencia. Las muestras se colocan en una celda transparente, que suele ser rectangular con un ancho de 1 cm. Las mejores celdas están hechas con cuarzo de alta calidad, aunque son comunes las de vidrio o plástico. El cristal, el vidrio y la mayoría de los plásticos absorben en el UV, lo que limita su utilidad a las longitudes de onda en el visible.

#### 4.3.2. Espectrofotómetro convencional

Un espectrofotómetro convencional enfoca la luz policromática de la fuente en un monocromador. Éste tiene como componentes principales una ranura de entrada, un elemento que dispersa la luz en su composición por longitudes de onda y una ranura de

salida que permite seleccionar la longitud de onda deseada. Esa luz monocromática atraviesa la muestra y llega al detector (Fig. 4.2). Las mediciones fotométricas se hacen en base a la relación entre la intensidad de la luz que alcanza al detector cuando está interpuesta la muestra y cuando no lo está.

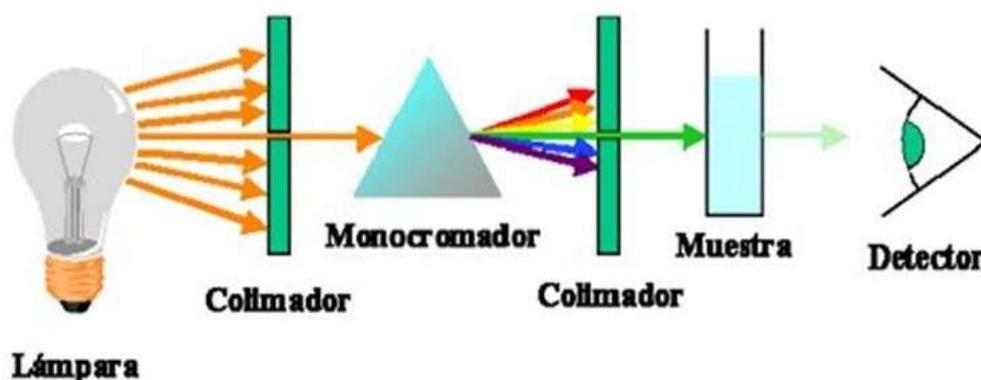


Fig. 4.2: Diagrama esquemático de un espectrofotómetro UV-Vis.

En realidad, el monocromador no selecciona una única longitud de onda, sino un rango, cuya amplitud depende de la calidad del mismo. Esta resolución depende fundamentalmente del diseño del monocromador, de su distancia focal y de las dimensiones y densidad de líneas en la rejilla de difracción.

Para cambiar la longitud de onda de medición, o para hacer un barrido espectral, se mueve el elemento dispersor o algún espejo por medio de un motor por pasos.

La estructura y funcionamiento general del espectrofotómetro convencional servirá como base tanto para el arreglo experimental, que se describe en la sección de resultados, como para la selección de datos en la implementación de la red neuronal descrita en el capítulo siguiente.

# Capítulo 5

## Redes neuronales

El enfoque convencional de la informática se basa en un conjunto explícito de instrucciones programadas. Las redes neurales o neuronales representan un paradigma computacional alternativo en que la solución de un problema se aprende de un conjunto de ejemplos. La inspiración para las redes neuronales se origina de los estudios de los mecanismos de procesamiento de la información en los sistemas nerviosos biológicos, en particular el cerebro humano. De hecho, gran parte de la investigación actual sobre algoritmos de redes neuronales se centra en obtener una comprensión más profunda de procesamiento de la información en los sistemas biológicos.

Una red neuronal *feedforward* puede considerarse como una función matemática no lineal que transforma un conjunto de variables de entrada en un conjunto de variables de salida. La forma exacta de la transformación se rige por un conjunto de parámetros llamados pesos cuyo valor puede determinarse sobre la base de un conjunto de ejemplos de la correlación necesaria. El proceso de determinación de estos valores de los parámetros a menudo se llama aprendizaje o formación, y puede ser una tarea computacionalmente intensa. Una vez que los pesos se han fijado, sin embargo, los nuevos datos pueden ser procesados por

la red muy rápidamente. Resulta conveniente en varios puntos de esta revisión establecer una analogía entre las redes neuronales artificiales y la técnica estándar de ajuste de curvas usando funciones.

## 5.1. Redes neuronales biológicas

El cerebro humano es la estructura más compleja conocida y la comprensión de su funcionamiento representa uno de los retos más difíciles y emocionantes que se enfrentan por la ciencia. Las redes neuronales biológicas proporcionan una fuerza impulsora detrás de una gran cantidad de investigación en modelos de redes artificiales, que es complementaria a la voluntad de construir mejores sistemas de reconocimiento de patrones y procesamiento de información. Para completar se proporciona esquema simplificado de las redes neuronales biológicas.

El cerebro humano contiene alrededor de  $10^{11}$  células eléctricamente activas llamadas neuronas. Éstas existen en una gran variedad de formas, aunque la mayoría tienen las características comunes que se indican en la Fig. 5.1. El árbol de ramificación de las dendritas proporciona un conjunto de entradas a la neurona, mientras que el axón actúa como una salida. La comunicación entre las neuronas se lleva a cabo en las uniones llamadas sinapsis. Cada neurona típicamente hace las conexiones a muchas miles de otras neuronas, por lo que el número total de las sinapsis en el cerebro excede  $10^{14}$ . A pesar de que cada neurona es un sistema de procesamiento de información relativamente lento (que opera en una escala de tiempo efectivo de alrededor de 1 ms) el paralelismo masivo de procesamiento de la información en muchas sinapsis conduce simultáneamente a una potencia de procesamiento eficaz, que es muy superior a la de las supercomputadoras actuales. También conduce a un alto grado de

tolerancia a fallos, como ejemplo muchas neuronas mueren cada día con poco efecto adverso en el rendimiento.

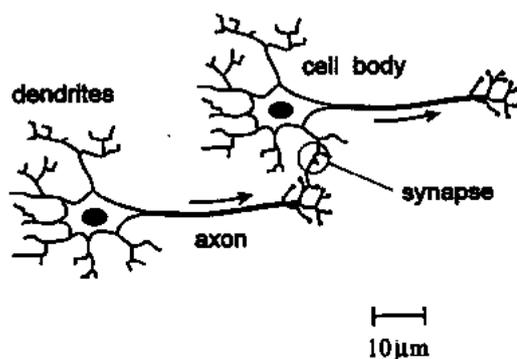


Fig. 5.1: Ilustración esquemática de dos neuronas biológicas. Imagen tomada de [13].

Muchas neuronas actúan de una manera todo-o-nada, una señal se envía por medio de un impulso eléctrico (llamado un potencial de acción) que se propaga desde el cuerpo celular a lo largo del axón, cuando esta señal llega a una sinapsis desencadena la liberación de neuro-transmisores químicos que atraviesan la unión sináptica a la siguiente neurona, dependiendo del tipo de sinapsis, esto puede aumentar (sinapsis excitatoria) o disminuir (sinapsis inhibitoria) la probabilidad de encender las siguientes neuronas. Cada sinapsis tiene una resistencia asociada (o peso) que determina la magnitud del efecto de un impulso en la neurona postsináptica. Cada neurona calcula de este modo una suma ponderada de las entradas de otras neuronas, y, si esta estimulación total supera cierto límite, las neuronas se encienden.

## 5.2. Redes neuronales artificiales

Un modelo matemático simple de una sola neurona se introdujo en un artículo por McCulloch y Pitts en 1943[13], y adopta la

forma indicada en la Fig. 5.2. Se puede considerar como una función no lineal, que transforma un conjunto de variables de entrada  $x_i$ , ( $i = 1, \dots, d$ ) en una variable de salida  $z$ . Se debe tener en cuenta que a partir de ahora se referirá a un modelo artificial de una neurona como una unidad de procesamiento, o simplemente unidad, para distinguirlo de su contraparte biológica. En el modelo de McCulloch-Pitts, la señal  $x_i$  en la entrada  $i$  se multiplica primero por un parámetro  $w_i$  conocido como un peso (que es análogo a la fuerza sináptica en una red biológica) y después se añade a todas las otras señales de entrada ponderada para dar un total de entrada a la unidad de la forma

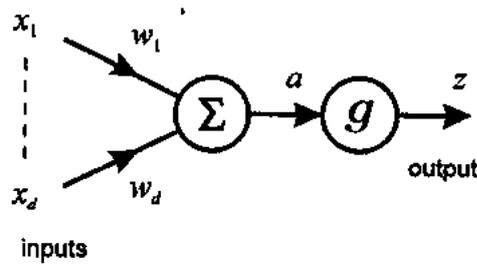


Fig. 5.2: Modelo de McCulloch-Pitts para una neurona simple con  $d$  entradas. Imagen tomada de [13].

$$a = \sum_{i=1}^d w_i x_i + w_0, \quad (5.1)$$

donde el parámetro  $w_0$  se llama sesgo o bias (y corresponde al límite de agotación en una neurona biológica). Formalmente, el sesgo puede ser considerado como un caso especial de un peso a partir de una entrada adicional cuyo valor ( $x_0$ ) se establece de forma permanente igual a  $x_0 = 1$ . Por lo tanto, podemos escribir la Ec. 5.1 en la forma

$$a = \sum_{i=0}^d w_i x_i. \quad (5.2)$$

Obsérvese que los pesos (y el sesgo) pueden ser de cualquier signo, que corresponde a sinapsis excitatorias o inhibitorias.

La salida de la unidad,  $z$ , (que vagamente puede considerarse como análoga a la tasa promedio agotador de una neurona) se da entonces mediante la operación de una función de activación no lineal  $g$  de tal forma que

$$z = g(a). \quad (5.3)$$

Algunas formas posibles para la función  $g$  se muestran en la Fig. 5.3.

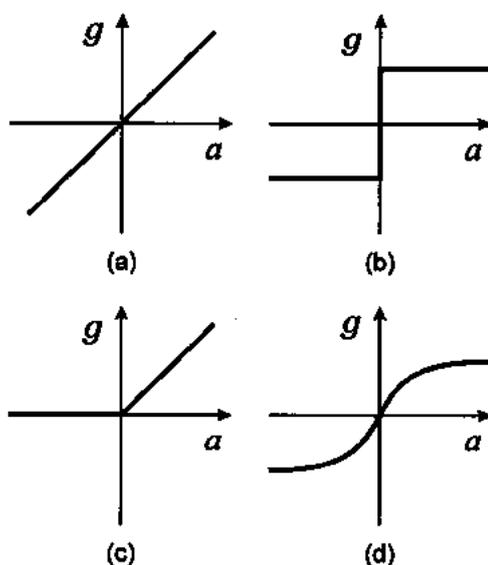


Fig. 5.3: Funciones típicas de activación: (a) Línea recta  $g(a) = a$ . (b) Escalón  $g(a) = 1 \forall a > 0$ ,  $g(a) = -1 \forall a < 0$ . (c) Línea acotada  $g(a) = a \forall a \geq 0$ . (d) Función sigmoide  $g(a) = \frac{1}{1+\exp^{-a}} - \frac{1}{2}$ . Imagen tomada de [13].

Aunque se ha introducido este modelo matemático de la neurona como una representación del comportamiento de neuronas biológicas, precisamente las mismas ideas también surgen cuando se consideran enfoques óptimos para la solución de problemas en el reconocimiento de patrones estadísticos. En este contexto, las expresiones tales como las ecuaciones 5.2 y 5.3 son conocidos como discriminantes lineales.

### 5.3. Función de error y formación de las redes

El problema de determinar los valores de los pesos en una red neuronal se llama entrenamiento y se introduce más fácilmente usando la analogía de ajuste de una curva polinomial a través de un conjunto de puntos de  $n$  datos. Se va a marcar un punto de datos en particular con el índice  $q = 1, \dots, n$ . Cada punto de datos consiste en un valor de  $x$ , denotado por  $x^q$ , y el correspondiente valor deseado para la salida  $y$ , que se designará por  $t^q$ . Estos valores de salida deseados se llaman valores objetivo en el contexto de la red neuronal. Se debe tener en cuenta que los puntos de datos a veces también se refieren como patrones. Con el fin de encontrar los valores adecuados para los coeficientes en el polinomio, es conveniente considerar el error entre el valor de salida deseado  $t^q$ , para una entrada en particular  $x^q$ , y el correspondiente valor predicho por la función polinomial dada por  $y(x^q; W)$ , donde  $W$  representa las distintas combinaciones de  $w_i$  que intervienen en el valor de  $y$ . Los procedimientos de ajuste de curvas estándar son los de minimizar el cuadrado de este error, sumado sobre todos los puntos de datos, dado por

$$E = \frac{1}{2} \sum_{q=1}^n [y(x^q; W) - t^q]^2. \quad (5.4)$$

Se considera a  $E$  como una función de  $W$ , por lo que la curva puede ser ajustada a los datos eligiendo un valor de  $W$  que minimiza  $E$ . Esto significa que el mínimo de  $E$  se puede encontrar en términos de la solución de un conjunto de ecuaciones algebraicas lineales.

La Fig. 5.4 muestra un ejemplo de un conjunto de puntos de datos junto con un polinomio cúbico que ha sido ajustado a los datos mediante la minimización del error de la suma de cuadrados. Vemos que la curva mínima de errores captura con éxito la

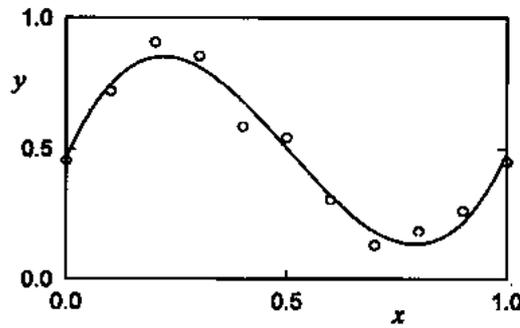


Fig. 5.4: Ejemplo del funcionamiento del ajuste cuadrático. Imagen tomada de [13].

tendencia subyacente en los datos.

La formación de una red neuronal procede de una manera análoga. Una función de error adecuada se define con respecto a un conjunto de puntos de datos, y los parámetros (pesos) se eligen para minimizar el error. Las funciones de las redes neuronales dependen de forma no lineal en sus pesos, por lo que la reducción al mínimo de la función de error correspondiente es sustancialmente más difícil que en el caso de polinomios, y generalmente requiere el uso de algoritmos de optimización no lineales iterativos.

En el caso de una red neuronal, cada vector de entrada  $x^q = (x_1^q, \dots, x_d^q)$  a partir del conjunto de datos tiene un correspondiente vector objetivo  $t^q$ . El error de la salida  $k$  cuando la red se presenta con el modelo,  $q$  viene dado por  $y_k(x^q; w) - t_k^q$ . El error total para todo el conjunto patrón entonces se puede definir como los cuadrados de los errores individuales sumado sobre todas las unidades de salida y sobre todos los patrones. Esto da

una función de error, de la forma

$$E = \frac{1}{2} \sum_{q=1}^n \sum_{k=1}^c [y_k(x^q; W) - t_k^q]^2. \quad (5.5)$$

## 5.4. Perceptrón multicapa

Hasta ahora se han descrito redes neuronales *feedforward* en términos de correlaciones no lineales entre espacios multidimensionales. Ahora se introducirá una forma explícita para el mapeo conocida como la red perceptrón multicapa. Esta clase de redes se han utilizado como base para la mayoría de las aplicaciones prácticas de redes neuronales hasta la fecha. En los temas pasados se introdujo el concepto de una sola unidad de procesamiento descrito por las Ecs. 5.2 y 5.3. Si se considera un conjunto de  $m$  unidades, todas con entradas comunes, entonces se llega a una red neuronal que tiene una sola capa de parámetros adaptativos (pesos) como se ilustra en la Fig. 5.5. Las variables de salida se indican mediante  $z_j$  y están dadas por

$$z_j = g(\sum_{i=0}^d w_{ji} x_i), \quad (5.6)$$

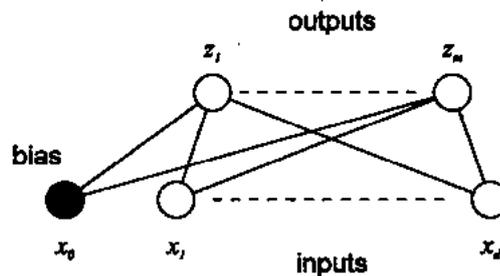


Fig. 5.5: Arreglo de la red simple donde se tienen  $d$  entradas y  $m$  salidas. Imagen tomada de [13].

donde  $w_{ji}$  es el peso de la entrada  $i$  a la unidad  $j$ , y  $g$  es una

función de activación como se discutió previamente.

La Fig. 5.6 muestra una red con dos sucesivas capas de unidades y por lo tanto de dos capas de pesos. Las unidades en la capa media se conocen como unidades ocultas ya que sus valores de activación no son directamente accesibles desde fuera de la red. La activación de estas unidades está de nuevo dada por la Ec. 5.6 como en el caso de la red de una sola capa. Las salidas de la red se obtienen actuando sobre la  $z$  de la segunda transformación, que corresponde a una segunda capa de unidades, para dar

$$y_k = \widehat{g}(\sum_{j=0}^m \widehat{w}_{kj} x_j), \quad (5.7)$$

donde  $\widehat{w}_{kj}$  denota un peso en la segunda capa de unión unidad oculta  $j$  a la unidad de salida  $k$ . Note que se ha introducido una unidad oculta extra con  $z_0 = 1$  para proporcionar un sesgo para las unidades de salida. Los sesgos (tanto para las unidades ocultas y de salida) desempeñan un papel importante en asegurar que la red puede representar mapeos no lineales generales. Podemos combinar las Ecs. 5.6 y 5.7 para dar la expresión completa para la transformación representada por la red en la forma

$$y_k = \widehat{g}(\sum_{j=0}^m \widehat{w}_{kj} g(\sum_{i=0}^d w_{ji} x_i)). \quad (5.8)$$

Una vez más, cada uno de los componentes de la ecuación 5.8 corresponde a un elemento del diagrama de la Fig. 5.6. Observe que la función de activación  $\widehat{g}$  aplicada a las unidades de salida no tiene que ser la misma que la función de activación  $g$  utilizado para las unidades ocultas.

Una opción común, y que se utilizará para ambas funciones de activación en esta tesis, para la función de activación es la sigmoide logística, representada en la Fig. fig:neurona3d sumando

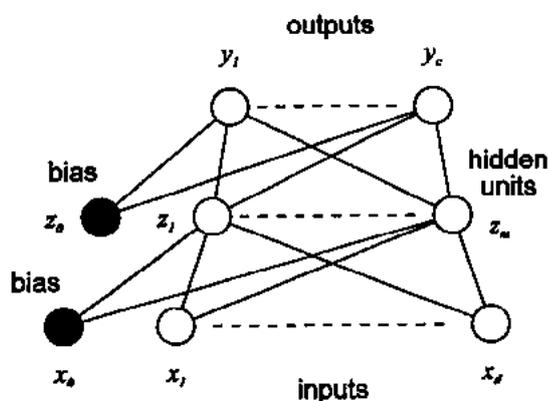


Fig. 5.6: Una red neuronal perceptrón multicapa que tiene dos capas de pesos. Imagen tomada de [13].

1/2 en la función, dada por

$$g(a) = \frac{1}{1 + \exp^{-a}}. \quad (5.9)$$

## 5.5. Método de entrenamiento *Backpropagation*

Conceptualmente, una red *feedforward* propaga hacia adelante una activación para producir una salida y hacia atrás se propaga el error para determinar los cambios que se requiere hacer a los pesos (como se muestra en la Fig. 5.7). Los pesos de las conexiones entre las neuronas median los valores pasados en ambas direcciones.

El algoritmo *backpropagation* (o propagación hacia atrás) se utiliza para actualizar los pesos de una red neuronal multicapa con una arquitectura fija. Se utiliza el método de descenso de gradiente para tratar de minimizar el error cuadrático entre los valores de salida de la red y los valores objetivos propuestos.

La Fig. 5.8 representa las componentes de una red que afectan a

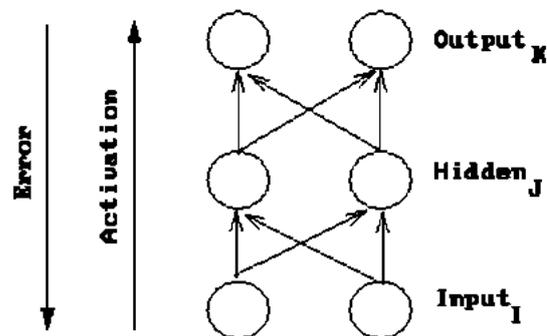


Fig. 5.7: Procedimiento en una red neuronal. Imagen tomada de [13].

un cambio de peso en particular. Obsérvese que todos los componentes necesarios están relacionados a nivel local con el peso involucrado. Ésta es una característica de la propagación hacia atrás que parece biológicamente plausible. Sin embargo, las conexiones del cerebro parecen ser unidireccionales y bidireccionales.

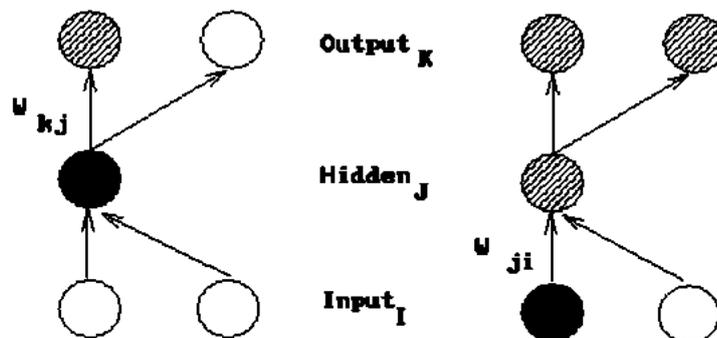


Fig. 5.8: Izquierda: El cambio de un peso oculto a la salida depende del error en el nodo de salida y la activación en el nodo oculto. Derecha: Mientras que el cambio a una entrada de peso oculto depende de error en la activación en el nodo oculto ( que a su vez depende del error en todos los nodos de salida) y en el nodo de entrada. Imagen tomada de [13].

### 5.5.1. Notación

Se construirá un algoritmo el cual actualizará los valores de los pesos en función de la variación que tiene el error para un peso específico, para ello se derivará el error con respecto el peso en cuestión, procedimiento que se explica en la siguiente sección. Para el propósito de esta derivación, utilizaremos la siguiente notación:

- El subíndice  $k$  denota la capa de salida.
- El subíndice  $j$  denota la capa oculta.
- El subíndice  $i$  denota la capa de entrada.
- $w_{kj}$  denota el peso de la capa oculta a la capa de salida.
- $w_{ji}$  denota el peso de la capa entrada a la capa oculta.
- $a$  denota un valor de activación.
- $t$  denota un valor objetivo.
- $net$  denota la entrada Neta.

### 5.5.2. Descenso de gradiente sobre el error

Podemos motivar al algoritmo de aprendizaje de propagación hacia atrás como descenso de gradiente sobre el error de la suma al cuadrado (elevamos al cuadrado el error, porque estamos interesados en su magnitud, no su signo). El error total en una red está dada por la siguiente ecuación.

$$E = \frac{1}{2} \sum_k (t_k - a_k)^2. \quad (5.10)$$

Se pretende ajustar los pesos de la red para reducir el error total

$$\Delta W \propto -\frac{\partial E}{\partial W}. \quad (5.11)$$

Empezando en la capa de salida con un peso determinado.

$$\Delta w_{kj} \propto -\frac{\partial E}{\partial w_{kj}}. \quad (5.12)$$

Como el error no es directamente una función del peso, se expande de la siguiente manera.

$$\Delta w_{kj} = -\epsilon \frac{\partial E}{\partial a_k} \frac{\partial a_k}{\partial net_k} \frac{\partial net_k}{\partial w_{kj}}. \quad (5.13)$$

Derivando el error con respecto la activación se obtiene

$$\frac{\partial E}{\partial a_k} = \frac{\partial}{\partial a_k} \left( \frac{1}{2} (t_k - a_k)^2 \right) = -(t_k - a_k). \quad (5.14)$$

Se sigue con la derivada del valor de activación con respecto la entrada neta en  $k$

$$\frac{\partial a_k}{\partial net_k} = \frac{\partial}{\partial net_k} (1 + e^{-net_k})^{-1} = \frac{e^{-net_k}}{(1 + e^{-net_k})^2}. \quad (5.15)$$

Si se desea expresar la Ec. 5.15 en términos de la función de activación se obtiene

$$\frac{e^{-net_k}}{(1 + e^{-net_k})^2} = \frac{e^{-net_k}}{1 + e^{-net_k}} = a_k(1 - a_k). \quad (5.16)$$

Se sigue con la derivación de la entrada neta en  $k$  con respecto el peso en cuestión. Notese que únicamente un término de la suma sobre la entrada neta tiene derivada, con respecto a un peso fijo, distinta de cero

$$\frac{\partial net_k}{\partial w_{kj}} = \frac{\partial w_{kj} a_j}{\partial w_{kj}} = a_j. \quad (5.17)$$

Sustituyendo las Ecs. 5.14, 5.16 y 5.17 en la Ec. 5.13 se obtiene

$$\Delta w_{kj} = \epsilon (t_k - a_k) a_k (1 - a_k) a_j. \quad (5.18)$$

Esta ecuación se simplifica típicamente como se muestra a continuación, donde el término  $\delta_k = (t_k - a_k)a_k(1 - a_k)$  representa el producto del error con la derivada de la función de activación.

$$\Delta w_{kj} = \epsilon \delta_k a_j. \quad (5.19)$$

Ahora se determina el cambio de peso apropiado para una entrada al peso de una neurona oculta. Esto es más complicado, ya que depende del error de todos los nodos que salen de esta conexión, la regla de cambio de peso para una entrada al peso de una oculta es dada por

$$\Delta w_{ij} \propto - \left[ \sum_k \frac{\partial E}{\partial a_k} \frac{\partial a_k}{\partial net_k} \frac{\partial net_k}{\partial a_j} \right] \frac{\partial a_j}{\partial net_j} \frac{\partial net_j}{\partial w_{ji}} \quad (5.20)$$

$$= \epsilon [\sum_k (t_k - a_k) a_k (1 - a_k) w_{kj}] a_j (1 - a_j) a_i. \quad (5.21)$$

Para simplificar la ecuación, si se define  $\delta_k = (t_k - a_k) a_k (1 - a_k)$ , se puede reescribir a la Ec. 5.21 como

$$\Delta w_{ij} = \epsilon [\sum_k \delta_k w_{kj}] a_j (1 - a_j) a_i. \quad (5.22)$$

Con el mismo argumento se llama  $\delta_j = [\sum_k \delta_k w_{kj}] a_j (1 - a_j)$ , obteniendo finalmente

$$\Delta w_{ij} = \epsilon \delta_j a_i. \quad (5.23)$$

## 5.6. Construcción del algoritmo

Para el desarrollo de la red neuronal, se utilizará una sola capa de neuronas ocultas, es necesario obtener un grupo de muestras, una relación par entrada - valor de salida deseado, las cuales

serán asociadas al patrón de entradas de la red neuronal. Este grupo de muestras se clasificarán a su vez en tres subgrupos, el primer subgrupo corresponde a los llamados valores de entrenamiento, el segundo subgrupo será llamado subgrupo de validación y el tercero será el subgrupo de predicción.

El funcionamiento de la red consiste en un aprendizaje de un conjunto predefinido de pares de entradas-salidas dados como ejemplo, empleando un ciclo propagación-adaptación de dos fases:

- 1 Propagación hacia adelante: Se aplica un patrón de entrada como estímulo para la primera capa de las neuronas de la red, se va propagando a través de todas las capas superiores hasta generar una salida, se compara el resultado obtenido en las neuronas de salida con la salida que se desea obtener y se calcula un valor del error para cada neurona de salida.
- 2 Propagación hacia atrás: Estos errores se transmiten hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de la capa intermedia que contribuyan directamente a la salida, recibiendo el porcentaje de error aproximado a la participación de la neurona intermedia en la salida original utilizando la regla de la cadena. Este proceso se repite, capa por capa, hasta que todas las neuronas de la red hayan recibido un error que describa su aportación relativa al error total. Basándose en el valor del error recibido, se reajustan los pesos de conexión de cada neurona, de manera que en la siguiente vez que se presente el mismo patrón, la salida esté más cercana a la deseada (disminuya el error).

El primer paso de la red, consiste en analizar los valores de entrenamiento por medio de la propagación hacia adelante. Definiendo  $a$ ,  $Net_k$  en la Fig. 5.9, de la Ec. 5.2, donde los primeros valores de los pesos son tomados de forma aleatoria para  $-1 < w_i < 1$

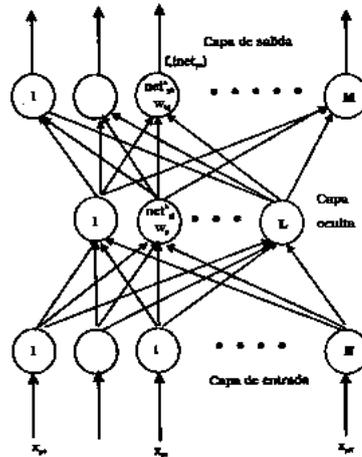


Fig. 5.9: Esquema de la relación entre las neuronas de entrada, ocultas y las de salida por medio de los pesos en una red neuronal. Imagen tomada de [14].

y los valores de  $x_i$  son los elementos que componen cada elemento del grupo de entrenamiento, los valores que tome  $a$  son asociados a las entradas de cada neurona oculta. La función de activación sobre cada neurona oculta, dada por la Ec. 5.9, es el valor obtenido al evaluar  $g(a)$  para toda neurona oculta, ahora este valor será asociado como  $x_k$  y funcionará como el término que acompaña a  $w_k$  en la Ec. 5.2 para obtener el valor de una nueva  $a$ ,  $Net_m$  en la Fig. 5.9, que serán los valores de entrada en la neurona de salida. De la misma forma que para las neuronas ocultas, la función de activación nos dará el valor de salida para cada neurona de la capa de salida. Por lo tanto, el conjunto de valores de resultantes de la capa de salida será el valor a comparar con el conjunto de valores de salida deseados del subgrupo de entrenamiento para obtener el error.

Para la segunda parte se sigue del error calculado en la primer parte y asociándolo con las Ecs. 5.19 y 5.23, que es la variación asociada a los pesos de las neuronas ocultas y de entrada respectivamente. La forma de construir la actualización de los valores de los pesos de las neuronas de entrada y ocultas están dadas

respectivamente por:

$$w_{ij}^n = w_{ij} + \Delta w_{ij} \quad (5.24)$$

y

$$w_{kj}^n = w_{kj} + \Delta w_{kj}. \quad (5.25)$$

Este procedimiento se repite para todos los elementos del subgrupo de entrenamiento. Una vez terminado este proceso se propaga únicamente hacia adelante con un elemento aleatorio del subgrupo de validación y calculamos el error sobre esta propagación, sin modificar los pesos y se guarda ese error. Repitiendo el proceso de entrenamiento y propagando hacia adelante hasta terminar los elementos del subgrupo de validación, se obtiene un error final conocido como el error de validación. Este error sirve para verificar que no se sobreentrene [14] la red. Repitiendo el procedimiento tantas veces como el número de elementos del subgrupo de validación, éste es una iteración en el ciclo completo en el algoritmo de entrenamiento de la red. Por construcción de la red, el error en la propagación siempre decrece, sin embargo esta condición no se cumple necesariamente en el error de validación, por lo tanto el algoritmo de entrenamiento se repite hasta que el error de entrenamiento sea más pequeño que un límite establecido a criterio del programador o el error de validación crezca, lo cual implicaría que la red se está sobreentrenando.

#### Algoritmo de entrenamiento

1. Inicializar los pesos con valores aleatorios
2. Presentar el patrón de entrada,  $X_p : x_{p1}, \dots, x_{pn}$ , y especificar la salida deseada :  $d_1, \dots, d_m$ .
3. Calcular la salida actual de la red :  $y_1, \dots, y_m$ . Para ello :
  - Se calculan las entradas netas para las neuronas ocultas procedentes de las neuronas de entrada

- Se calculan las salidas de las neuronas ocultas
  - Se realizan los mismos cálculos para obtener las salidas de las neuronas de salida
4. Calcular los términos de error para todas las neuronas
  5. Actualización de los pesos
  6. Repetición del proceso hasta terminar los elementos del subgrupo de validación, propagando hacia adelante en cada iteración un elemento de este subgrupo y calcular sus errores. Esto equivale a una iteración del ciclo completo de entrenamiento
  7. Repetición del ciclo hasta que el término de error resulte aceptablemente pequeño para cada uno de los patrones aprendidos o el error de validación crezca.

Finalmente teniendo los pesos calculados en la red previamente entrenada, se construye un algoritmo que propague únicamente hacia adelante con los pesos obtenidos en la última iteración del ciclo de entrenamiento, con la finalidad de propagar los elementos del subgrupo de predicción y ver que tan eficiente es la red.

# Capítulo 6

## Desarrollo experimental

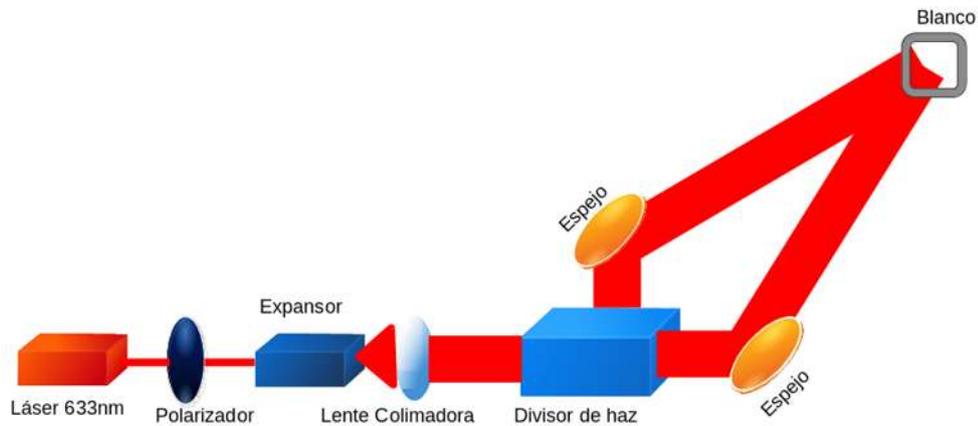
### 6.1. Construcción de la rejilla

El esquema experimental empleado para generar la rejilla holográfica se muestra en la Fig. 6.1a, el cual está constituido por un láser rojo de 633 nm, un polarizador con su eje de transmisión orientado verticalmente, un sistema colimador<sup>1</sup>, un divisor de haz, dos espejos colocados adecuadamente y una base para portaobjetos, donde se colocará posteriormente una película holográfica para grabar un patrón de interferencia. En la Fig. 6.1b se muestra una fotografía del arreglo experimental.

Los espejos se colocan a una distancia  $a$  entre ellos, de tal forma que las ondas planas reflejadas respectivamente por ellos incidan sobre una pantalla, situada en la base para el portaobjetos, colocada a una distancia  $s$  perpendicular al punto medio de la recta que une ambos espejos tal como se muestra en la Fig. 6.2, por lo que el patrón de interferencia corresponde a franjas cuya frecuencia obedece la Ec. 2.42.

---

<sup>1</sup>Un sistema colimador consiste en un expansor de haz, cuya función es generar una onda esférica, el cual es colocado a la distancia focal de una lente que produce una onda plana.



(a)



(b)

Fig. 6.1: (a) Esquema utilizado para la elaboración de la rejilla, (b) fotografía del arreglo.

Posteriormente se quitó la pantalla y se colocó un radiómetro para medir la irradiación generada por la superposición de ambas ondas planas en esta zona. El valor medido fue de  $5 \mu\text{W}/\text{cm}^2$ . La película holográfica que se utilizó es de la marca Slavich línea PFG 01 de  $75 \mu\text{J}/\text{cm}^2$  de sensibilidad (potencia) y como  $P = E/t$  se despeja el tiempo y da el valor  $t = 15 \text{ s}$ , donde  $t$  es el tiempo requerido para grabar el patrón de interferencia en la película holográfica.

Se construyeron cinco rejillas de distinta densidad de franjas<sup>2</sup>

<sup>2</sup>Se hicieron varias rejillas por la incertidumbre de la función pupila que delimita la

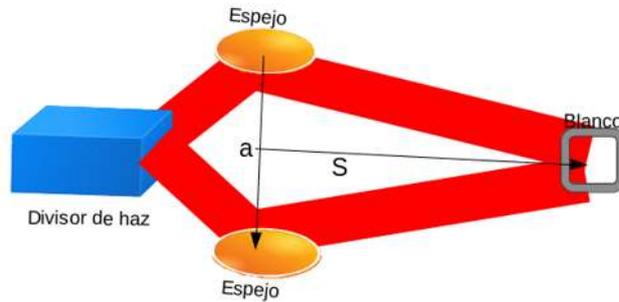


Fig. 6.2: Esquema de la posición de los espejos.

mostradas en la tabla 6.1. Una vez construidas se colocó la cámara del celular frente a cada una de las rejillas iluminadas con un LED para verificar si el espectro generado por las rejillas cae en su totalidad dentro de la función pupila de la cámara. Se observó que los espectros generados por cada una de las rejillas estaba dentro de los límites de la cámara, por lo que se eligió la que tenía un espectro más amplio que fue la correspondiente a  $\Delta y = 1266$ .

Tabla 6.1: Relación entre  $s$  y  $a$  con la distancia entre dos máximos en la rejilla.

$s$	$a$	$\Delta y$
35 cm	24 cm	923.125 nm
45 cm	30 cm	949.5 nm
35 cm	20 cm	1107.75 nm
24 cm	15 cm	1012.8 nm
24 cm	12 cm	1266 nm

---

cámara del celular.

## 6.2. Arreglo experimental

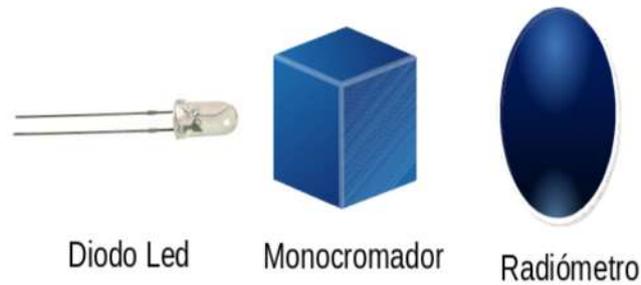
En el desarrollo del experimento se utilizó un LED comercial. Es de gran importancia saber la caracterización óptica del diodo, es decir la relación de la irradiancia respecto a cada longitud de onda que emite este LED. Esta relación fue obtenida con el arreglo descrito en la Fig. 6.3, el cual consiste en colocar el LED frente a un monocromador<sup>3</sup> seguido de un radiómetro. En este arreglo óptico se cambia la salida del monocromador y se mide la intensidad de salida con el radiómetro. Con este procedimiento se obtuvo una relación discreta mostrada en la Fig. 6.4 donde el eje  $y$  representa la medición de irradiancia en *nanoWatts* y el eje  $x$  son las longitudes de onda en el rango de 400 nm a 690 nm.

Otra limitante en las consideraciones de la imagen es la sensibilidad de la cámara del celular a una longitud de onda, es decir, no todo el espectro visible es captado por la cámara, ya sea por la misma sensibilidad de la cámara o por la irradiancia a la cual emite el LED. Este problema fue solucionado al sustituir el radiómetro en la Fig. 6.3 por el celular. Variando el monocromador y usando la cámara del celular se obtuvo que la sensibilidad de la cámara a este LED fue en el rango de 402 nm a 688 nm.

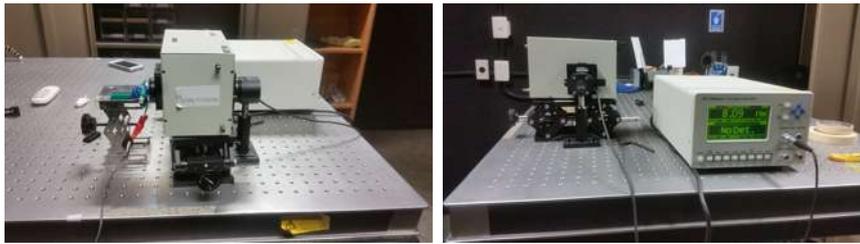
Teniendo las características del LED y de la cámara del celular, se montó un arreglo como se muestra en la Fig. 6.5, el cual consiste en un LED, una base para celdas, una rejilla de difracción y el celular colocado con  $\theta = 27^\circ$ , este ángulo fue seleccionado debido a las limitantes del arreglo, pero está en el rango de la Ec. 3.2. Para verlo nótese que se trabaja con el orden de difracción  $m = 1$  y que la luz incide aproximadamente normal a la rejilla,

---

<sup>3</sup>Un monocromador es un dispositivo óptico que deja pasar solo una longitud de onda de todo el espectro que constituye una fuente de luz, así mismo también permite variar esta longitud de onda emitida



(a)



(b)

(c)

Fig. 6.3: (a) Diagrama esquemático del arreglo, (b) vista lateral del arreglo (c) vista frontal del arreglo.

esto es  $\sin \alpha \approx 0$ , por lo que  $\sin \beta = \lambda/d \approx 1/2$  lo cual implica que  $\beta \approx 30^\circ$ , entonces el ángulo al que fue colocada la cámara del celular esta en un rango aceptable.

Las primeras imágenes que se tomaron fue el espectro sin muestra, seguida por una foto del espectro tomado a la muestra únicamente con agua, Fig. 6.6, estas dos serán utilizadas, en la siguiente sección, como punto de referencia en la comparación entre los espectros.

La sustancia a analizar fue un colorante verde natural de McCor-

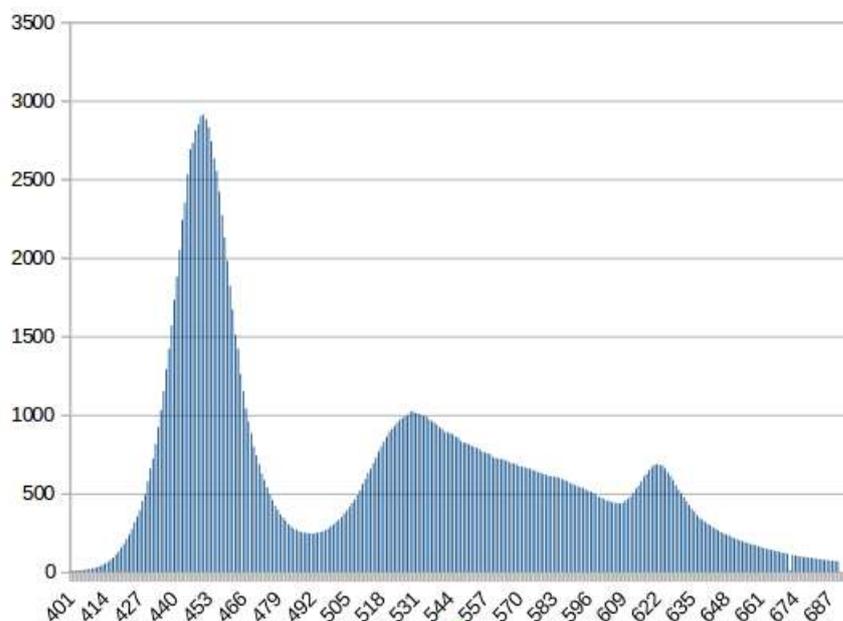
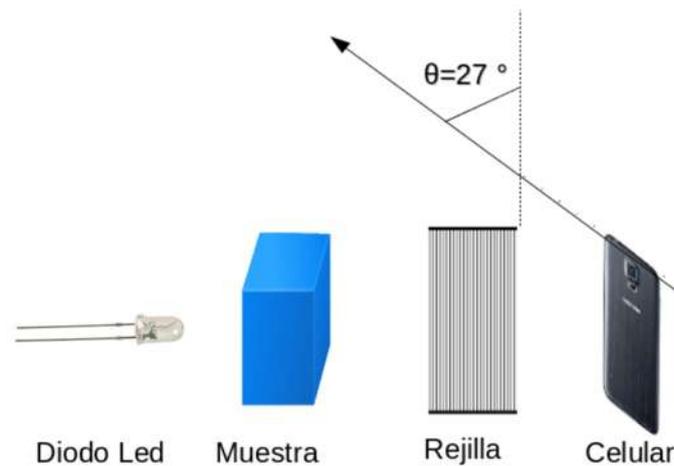


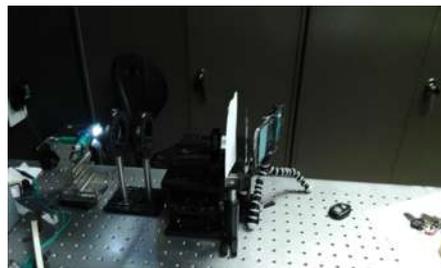
Fig. 6.4: Relación entre la irradiancia contra longitud de onda emitida por el LED  $I$  vs  $\lambda$ .

nic y el solvente fue agua. Las muestras se hicieron en relación de  $\frac{1}{16}$  ml equivalente a una porción de colorante por una porción constante de agua, fijada en 240 ml. Se tomaron 53 muestras empezando en 3 porciones de colorante por una de agua hasta llegar a 55 porciones de colorante por una de agua. Algunos espectros se muestran en la Fig. 6.7.

A su vez las distintas concentraciones fueron analizadas con el espectrofotómetro UV-Vis del laboratorio de óptica para ver el comportamiento de sus absorbancias. Algunas relaciones de la absorbancia respecto a la longitud de onda obtenidas por el espectrofotómetro son mostradas en la Fig. 6.8.



(a)



(b)

Fig. 6.5: (a) Esquema del arreglo utilizado para la obtención del espectro con el dispositivo móvil, (b) fotografía del arreglo.

### 6.3. Construcción de la red neuronal

La implementación de la red se desarrolló en la plataforma de MatLab. Para ello se utilizó el comando **imread** cuya función es reconocer una imagen y por default, MatLab asocia una imagen con tres matrices, RGB (Red, Green, Blue). Cada posición de la matriz corresponde a los valores de cada pixel de la imagen, cuya escala es de 1 a 255. Para simplificar el código se analizará una sola matriz cuyos valores son la suma de las tres matrices que representan a la imagen en cuestión.



Fig. 6.6: (a) Espectro tomado cuando no hay muestra, (b) espectro tomado a la muestra únicamente con agua.

Es importante saber que de la imagen a analizar solo importa la región donde se encuentra el espectro, por ello se asocia un vector al renglón que divide aproximadamente a la mitad la posición en la cual se encuentra el espectro, esto también ayuda a minimizar el tiempo de ejecución del algoritmo. Considérese la Fig. 6.6(b) (el espectro del agua), en la Fig. 6.9(a) se muestra el renglón asociado a la región media donde se encuentra el espectro. Ahora bien, este vector contiene datos de la imagen que no son de interés (los datos que no contienen información del espectro), por lo cual se puede reducir el tamaño del vector a la región únicamente donde se encuentra el espectro, en los datos que son distintos de 1, lo cual nos lleva al vector representado por la Fig. 6.9(b).

Como se hizo mención, en la espectroscopía la relación entre irradiancia y concentración de la sustancia es importante, por ellos se considera la caracterización del LED, descrita por la Fig. 6.4, esta relación se realizó de forma discreta, por lo que resulta impráctico tratar de analizar todo el vector mostrado en la Fig. 6.9(b) pues en general el tamaño del vector a analizar es de mayor tamaño que el vector asociado a la caracterización del LED. Una forma de resolver esto fue reduciendo el tamaño del vector al número de longitudes de onda caracterizadas del led, así se reduce el tamaño del vector de una forma conveniente. En este mismo procedimiento, la escala fue normalizada dos veces, la primera se hizo sobre el valor máximo que puede obtenerse de

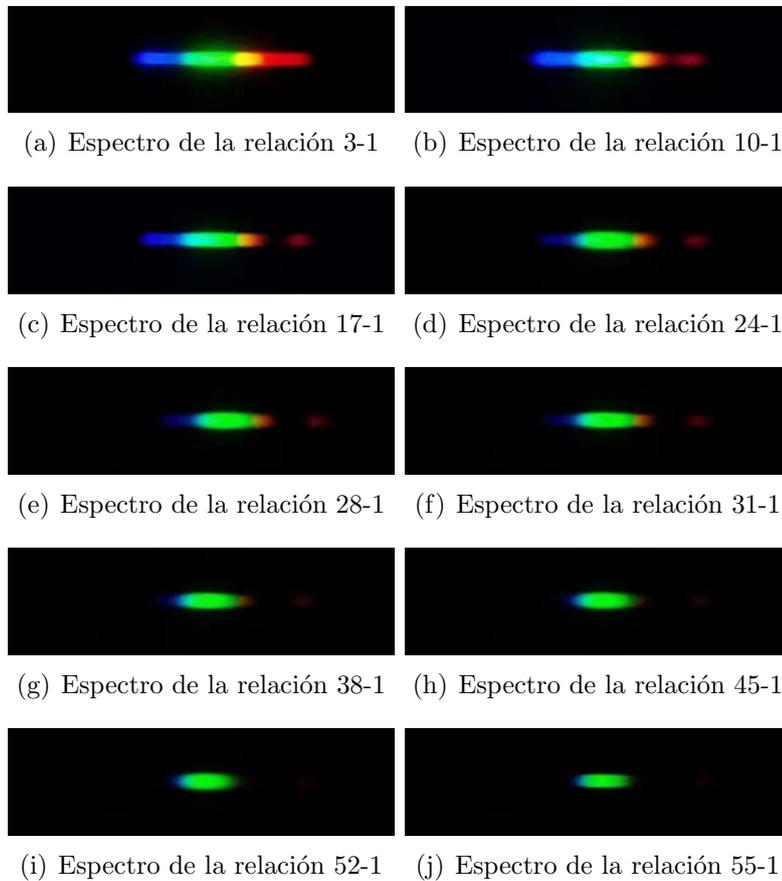


Fig. 6.7: Espectros de las muestras a distintas concentraciones.

un pixel, que es 255, la segunda normalización viene del hecho que la espectroscopía relaciona la irradiancia con la concentración de la sustancia. Esta implicación conlleva a normalizar el vector sobre los valores medidos por la caracterización óptica del led, en el intervalo de 402 nm a 688 nm que corresponde al rango donde la cámara del celular tiene sensibilidad al espectro del LED. El vector resultante es representado en la Fig. 6.9(c).

Para este caso en particular el tamaño del vector es de 287 elementos, por ello se construye una red en la cual la cantidad de neuronas de la capa oculta sea igual a la cantidad de elementos que tenga el vector de entrada. La cantidad de pesos asociados

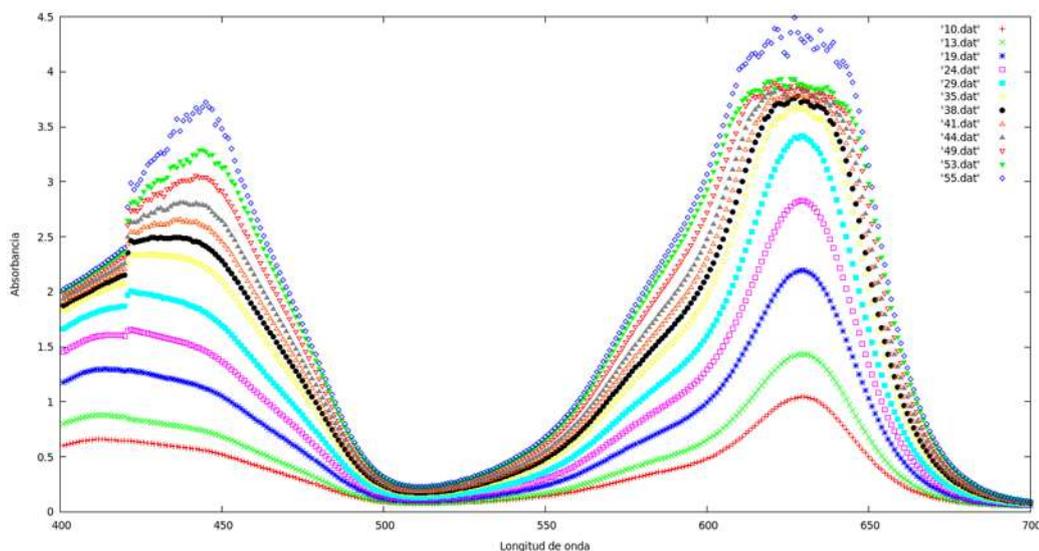


Fig. 6.8: Relación entre la absorbancia contra longitud de onda de las muestras.

al vector de entrada será de  $287 * 287 = 82369$ , estos pesos se deberán actualizar en cada iteración. Por lo tanto, como el algoritmo de entrenamiento se itera 33 veces por cada elemento del subgrupo de validación, estas son 330 iteraciones, el total de pesos calculados en una iteración del ciclo total de entrenamiento es de  $287 * 287 * 330 = 27,181,770$  pesos sólo para el conjunto de elementos de entrada por iteración, por este motivo se decidió acotar estos 287 elementos en bloques de 10. En total se redujo el vector de entrada de 287 elementos a 29, reduciendo la cantidad de actualizaciones de los pesos en 277,530 por iteración. El vector final es mostrado en la Fig. 6.9(d).

Para reducir el tamaño del vector en la forma descrita anteriormente se sigue el siguiente algoritmo:

1. Leer una imagen y asociar un vector al renglón a analizar.
2. Acotar el vector a la región de la imagen donde se encuentra el espectro.

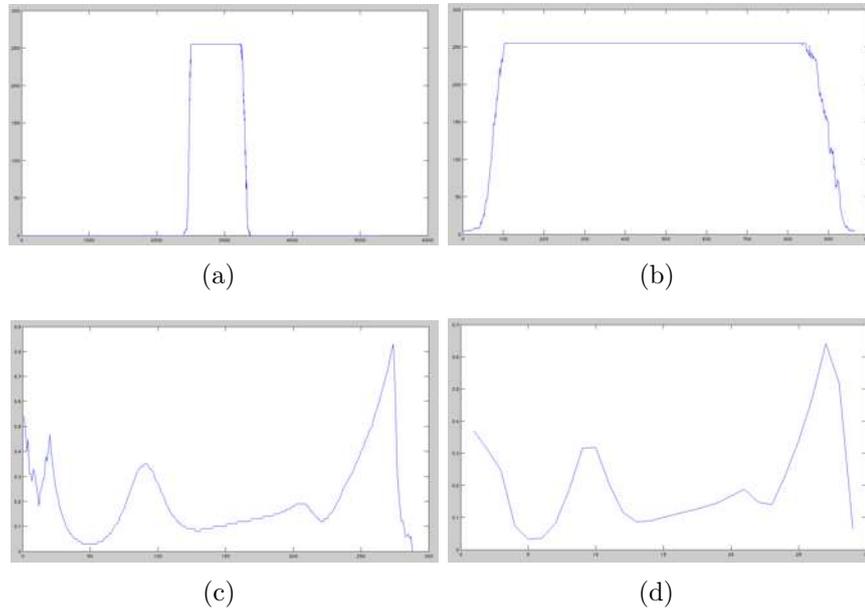


Fig. 6.9: (a) Vector de la imagen del espectro del agua, (b) vector en el intervalo donde está el espectro del agua, (c) vector acotado por la relación entre el tamaño del renglón de la región donde está el espectro y el tamaño del vector asociado a la caracterización del LED, normalizado por la caracterización óptica del LED y el valor máximo de un pixel, 255, (d) vector segmentado en bloques de 10.

3. Asociar la longitud del vector con el número de elementos que contenga la caracterización del LED.
4. Normalizar el vector dos veces, con respecto al valor máximo que puede tomar un pixel (que es 255) y con la caracterización del led.
5. Segmentar el vector normalizado por regiones específicas, en el caso de esta tesis fueron regiones de 10 elementos.

Como se menciona en el capítulo anterior de las 53 muestras tomadas se dividieron en 3 subgrupos. Para el caso del subgrupo de entrenamiento consiste en 33 elementos y los subgrupos de validación y predicción tienen 10 elementos cada uno. Cada elemento del subgrupo es un par de vectores, el primero asociado a

la entrada (vector relacionado a la comparación entre el espectro a analizar con el agua) y el segundo corresponde al valor de salida deseado. Para cada valor de salida deseado se construyó un vector de 33 renglones, cada posición representa que imagen de entrenamiento se está analizando. En el caso de los elementos de entrenamiento, si se analiza la imagen de la concentración 3-1 que corresponde al primer elemento del subgrupo de entrenamiento, entonces el vector de salida deseada se compone por un valor de 0.8 en su primer entrada y 0.2 de en las 32 restantes entradas, los vectores de salida deseada de todos los subgrupos toman valores en el intervalo 0.2 a 0.8<sup>4</sup>. De esta forma se puede ver el vector de salida en forma general como

$$Vec_i^{salida} = \sum_j^n (\delta_{ij} * 0,6 + 0,2)e_j, \quad (6.1)$$

donde  $\delta_{ij}$  es la delta de kronecker y  $e_j$  es el vector unitario de la  $j$ -ésima componente del vector de salida, para el caso del subgrupo de entrenamiento. Por otra parte, los elementos de los subgrupos de validación y predicción, se encuentran entre dos elementos del subgrupo de entrenamiento, por ello el vector de salida deseada para los elementos de estos subgrupos toma en cuenta la concentración del elemento en cuestión y los elementos del subgrupo de entrenamiento en que se encuentra. Por ejemplo, el primer elemento del subgrupo de validación es de relación 4-1, los primeros dos elementos del subgrupo de entrenamiento son 3-1 y 5-1, esto quiere decir que la concentración de este primer elemento de validación está entre el primer y segundo valor de entrenamiento, por ello se asocia el vector de salida a  $0,4e_1 + 0,4e_2 + 0,4e_3$  y 0,2 a todas las demás posiciones, donde

---

<sup>4</sup>La función de activación que se utilizó para programar la red, Ec. 5.9, limita los valores que puede arrojar la red entre 0 y 1, sin embargo estos valores se toman en el límite que  $a$  tienda a más o menos infinito, por ello se decide que los valores deseados esten en un rango de 0.2 a 0.8.

añadimos el elemento  $e_3$  debido al error al momento de hacer las muestras que es de  $\pm 1$  porciones de colorante. Teniendo este criterio y aplicándolo sobre todos los elementos de estos dos subgrupos se construyen los restantes vectores de salida deseada. Los elementos que componen cada subgrupo se muestran en la Fig. 6.10.

Entrenamiento		Validación		Predicción	
Ent.	Sal.+0,2e <sub>i</sub>	Ent.	Sal.+0,2e <sub>i</sub>	Ent.	Sal.+0,2e <sub>i</sub>
3-1	0,6e <sub>1</sub>	4-1	0,2(e <sub>1</sub> + e <sub>2</sub> + e <sub>3</sub> )	7-1	0,2(e <sub>3</sub> + e <sub>4</sub> + e <sub>5</sub> )
5-1	0,6e <sub>2</sub>	10-1	0,2(e <sub>5</sub> + e <sub>6</sub> + e <sub>7</sub> )	13-1	0,2(e <sub>7</sub> + e <sub>8</sub> + e <sub>9</sub> )
6-1	0,6e <sub>3</sub>	16-1	0,2(e <sub>9</sub> + e <sub>10</sub> + e <sub>8</sub> )	19-1	0,2(e <sub>11</sub> + e <sub>12</sub> + e <sub>13</sub> )
8-1	0,6e <sub>4</sub>	22-1	0,2(e <sub>13</sub> + e <sub>14</sub> + e <sub>15</sub> )	24-1	0,2(e <sub>14</sub> + e <sub>15</sub> + e <sub>16</sub> )
9-1	0,6e <sub>5</sub>	27-1	0,2(e <sub>16</sub> + e <sub>17</sub> + e <sub>18</sub> )	30-1	0,2(e <sub>18</sub> + e <sub>19</sub> + e <sub>20</sub> )
11-1	0,6e <sub>6</sub>	34-1	0,2(e <sub>21</sub> + e <sub>22</sub> + e <sub>23</sub> )	35-1	0,2(e <sub>21</sub> + e <sub>22</sub> + e <sub>23</sub> )
12-1	0,6e <sub>7</sub>	38-1	0,2(e <sub>23</sub> + e <sub>24</sub> + e <sub>25</sub> )	41-1	0,2(e <sub>25</sub> + e <sub>26</sub> + e <sub>27</sub> )
14-1	0,6e <sub>8</sub>	43-1	0,2(e <sub>26</sub> + e <sub>27</sub> + e <sub>28</sub> )	44-1	0,2(e <sub>26</sub> + e <sub>27</sub> + e <sub>28</sub> )
15-1	0,6e <sub>9</sub>	47-1	0,2(e <sub>28</sub> + e <sub>29</sub> + e <sub>30</sub> )	49-1	0,2(e <sub>29</sub> + e <sub>30</sub> + e <sub>31</sub> )
17-1	0,6e <sub>10</sub>	51-1	0,2(e <sub>30</sub> + e <sub>31</sub> + e <sub>32</sub> )	53-1	0,2(e <sub>31</sub> + e <sub>32</sub> + e <sub>33</sub> )
18-1	0,6e <sub>11</sub>				
20-1	0,6e <sub>12</sub>				
21-1	0,6e <sub>13</sub>				
23-1	0,6e <sub>14</sub>				
25-1	0,6e <sub>15</sub>				
26-1	0,6e <sub>16</sub>				
28-1	0,6e <sub>17</sub>				
29-1	0,6e <sub>18</sub>				
31-1	0,6e <sub>19</sub>				
32-1	0,6e <sub>20</sub>				
33-1	0,6e <sub>21</sub>				
36-1	0,6e <sub>22</sub>				
37-1	0,6e <sub>23</sub>				
39-1	0,6e <sub>24</sub>				
40-1	0,6e <sub>25</sub>				
42-1	0,6e <sub>26</sub>				
45-1	0,6e <sub>27</sub>				
46-1	0,6e <sub>28</sub>				
48-1	0,6e <sub>29</sub>				
50-1	0,6e <sub>30</sub>				
52-1	0,6e <sub>31</sub>				
54-1	0,6e <sub>32</sub>				
55-1	0,6e <sub>33</sub>				

Fig. 6.10: Relación de los elementos que conforman los subgrupos.

La forma en que funciona un espectrofotómetro convencional es comparar la irradiancia emitida por cada longitud de onda de una muestra en comparación de la irradiancia emitida por el solvente en ausencia del analito. Con el mismo algoritmo en que

se redujo el vector mostrado en la Fig. 6.9, se reducirá la comparación del espectro del agua con el espectro de las muestras de todos los subgrupos, para así obtener los vectores de entrada para la red. Los vectores que serán la entrada de la red neuronal, asociados con la comparación entre los espectros de las imágenes en la Fig. 6.7 con el espectro de la agua 6.6(b), son los mostrados en la Fig. 6.11. Nótese que las realciones mostradas en la Fig. anterior muestran un caracter similar al mostrado en la Fig. 6.8.

Una vez definidos los vectores con los que trabajará la red se sigue la construcción, la cual consiste básicamente en tres códigos, actualizar los pesos, propagación hacia delante y propagación hacia atrás, los códigos están en el apéndice A, B y C respectivamente. La función del código que actualiza los pesos es, recibiendo el vector de entrada ( $I$ ), vector de salida deseado ( $target$ ), los pesos predefinidos ( $w_i, w_o$ ) y el tamaño de paso ( $\epsilon$ ), propaga hacia adelante la red obteniendo el  $vector^{salida}$  como resultado de esta propagación y a su vez calcula su  $error$ . Posteriormente, tal como define las Ecs. 5.24 y 5.25, ajusta los pesos y los devuelve como los pesos actualizados  $wn_i$  y  $wn_o$ . Por su parte, la propagación hacia adelante recibe pesos fijos ( $w_i$  y  $w_o$ ), vector de entrada ( $I$ ) y el vector de salida deseada ( $target$ ) y devuelve el  $vector^{salida}$  junto con su respectivo  $error$  al propagar la red sin modificar los pesos de entrada. La propagación hacia atrás recibe un número máximo de iteraciones ( $iter$ ), el error máximo ( $errorMax$ ) y los subgrupos de entrenamiento y validación. Es importante conocer que este algoritmo, descrito en la sección 5.6, tiene 3 formas de detenerse la primera es si supera el límite de iteraciones puesta, la segunda si el error de entrenamiento es más pequeño que un límite puesto por  $errorMax$  y la última condición será si el error de validación crece más que algún punto que aún no se conoce.

Como en principio no se tiene información del comportamiento de la red neuronal al momento de la validación de datos, se corrió el programa cien mil veces para ver como era el comportamiento del error de validación, el resultado es mostrado en la Fig. 6.12.

La información que proporcionó la Fig. 6.12 confirma que el error de entrenamiento disminuye al aumentar las iteraciones lo cual implica que la red esta bien programada. También se puede observar que el valor mínimo que obtiene el error de validación es de 0.0343 y este es alcanzado entre las mil y dos mil iteraciones. Con esta información se modifica el *errorMax* en el mismo algoritmo para que se detenga si el error de validación está por debajo de 0.035. Con esta condición se obtuvieron las gráficas mostradas en la Fig. 6.13.

El error máximo se alcanzó en un total de 1365 iteraciones. Del algoritmo de propagación hacia atrás se obtienen los valores de los pesos  $w_i$  y  $w_o$  que representan los pesos asociados a las neuronas de entrada u ocultas respectivamente. Estos pesos se mantendrán fijos en el algoritmo de propagación hacia adelante para propagar los elementos del subgrupo de predicción y ver que tan eficiente es la red. Al propagar los elementos del subgrupo de predicción en la red se obtienen los vectores de salida mostrados en la Fig. 6.14.

Los vectores de salida deseada fueron construidos de esta forma para representar la imagen a la que más se parece el espectro, según sea la región donde se encuentre el máximo. Como los elementos del subgrupo de predicción se encuentran entre dos elementos del subgrupo de entrenamiento, la información que nos interesa del vector de salida de la red es una región en la cual se encuentre algún máximo. Esta región debe tomar en cuenta el error con la que se hicieron las muestras que es de  $\pm 1$  porción de analito. Por ello se tomó el promedio sobre 4

elementos del vector de salida y el valor del promedio más grande comprenderá la región en la que se conciderá más parecida la muestra del analito. El algoritmo que toma la posición del valor máximo del promedio sobre el vector de salida y lo compara con el vector de salida deseada se describe en el apéndice D llamado *AnalisisRedFinal*. Ingresando los vectores de salida de la red y los vectores de salida desada al algoritmo de comparación arrojó los resultados vistos en la Fig. 6.15.

Al ingresar los elementos del subgrupo de predicción con el algoritmo *AnalisisRedFinal* se obtuvo que la red predijo la región para 8 de los 10 elementos de dicho subgrupo, por lo cual se afirma que la red tiene una efectividad de al menos 80 %.

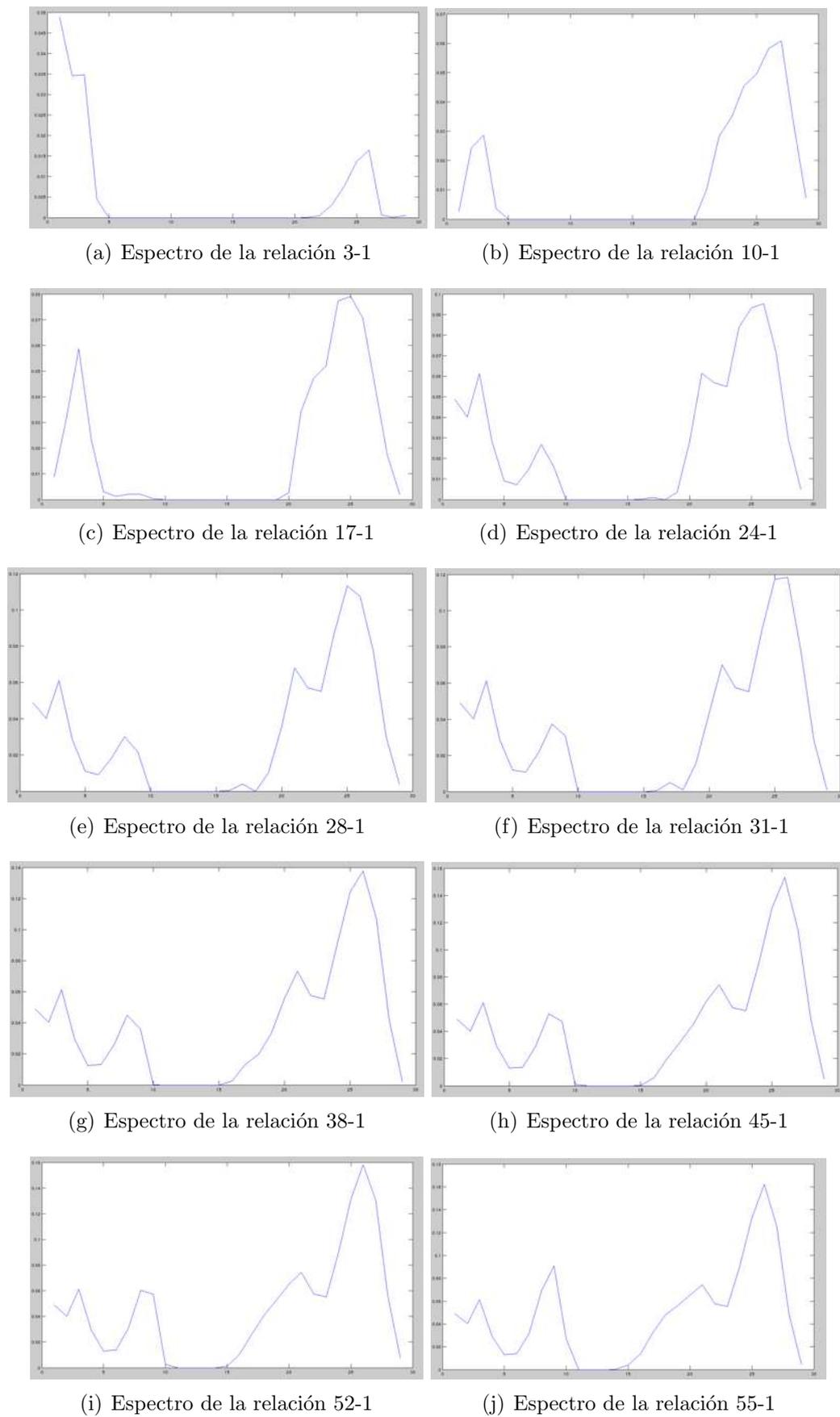
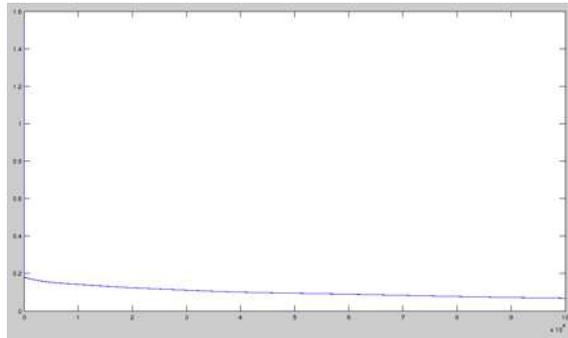
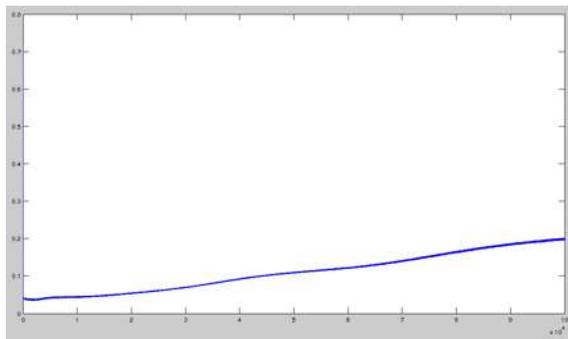


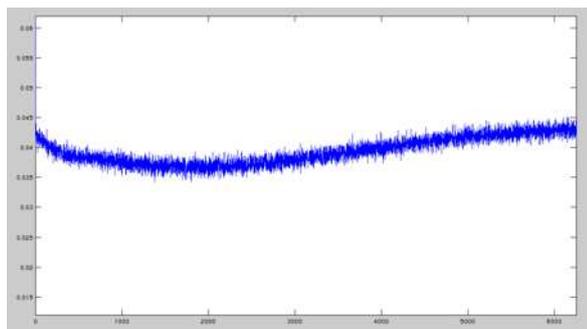
Fig. 6.11: Espectros de las muestras a distintas concentraciones.



(a)

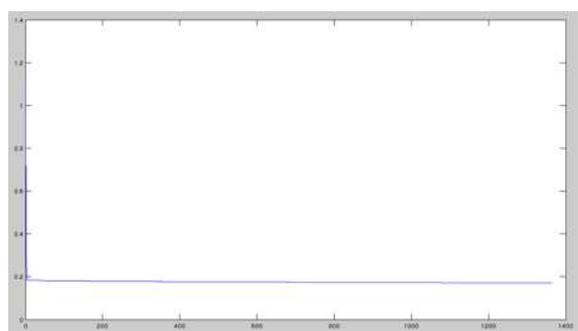


(b)

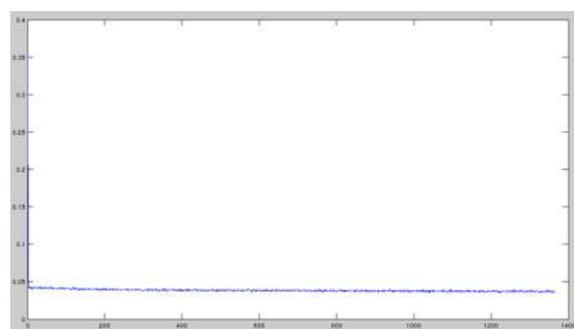


(c)

Fig. 6.12: (a) Error sobre el entrenamiento, (b) error de validación, (c) error de validación en el rango de 0 a seis mil iteraciones.

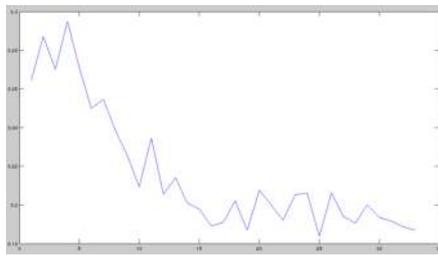


(a)

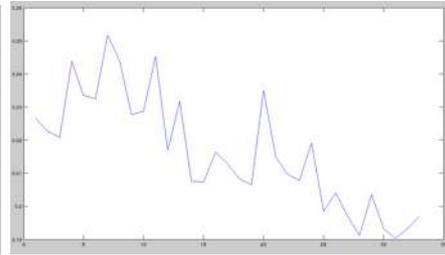


(b)

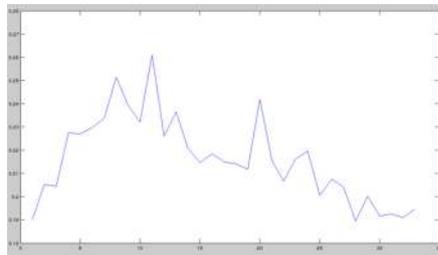
Fig. 6.13: Error de entrenamiento y de validación de la red neuronal fijando el error máximo de validación en 0.035. (a) Error sobre el entrenamiento, (b) error de validación.



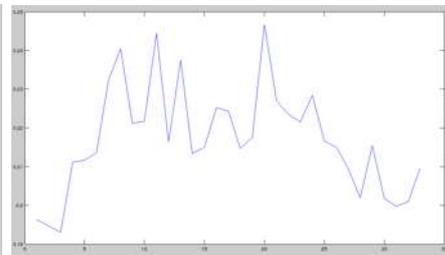
(a) Vector de salida del elemento 1 del subgrupo de validación



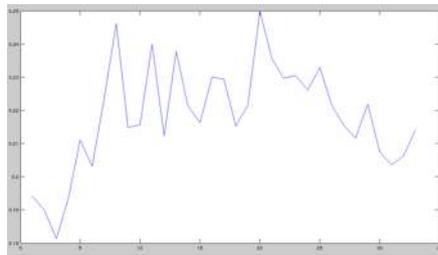
(b) Vector de salida del elemento 2 del subgrupo de validación



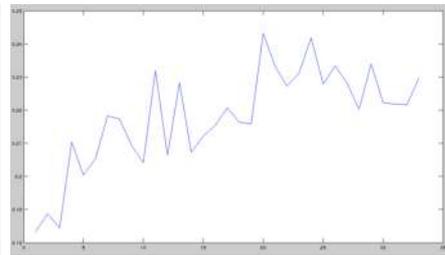
(c) Vector de salida del elemento 3 del subgrupo de validación



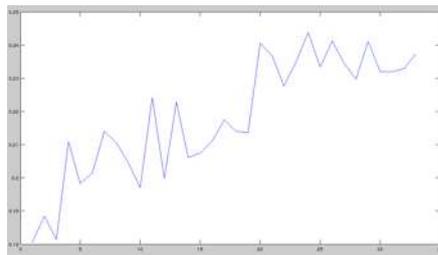
(d) Vector de salida del elemento 4 del subgrupo de validación



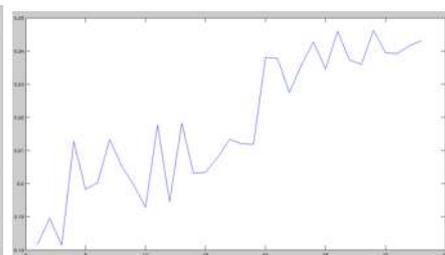
(e) Vector de salida del elemento 5 del subgrupo de validación



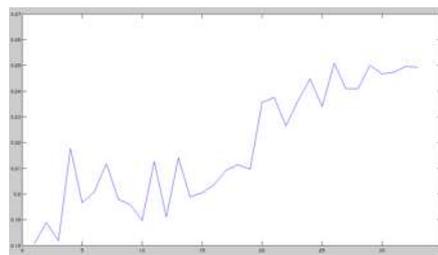
(f) Vector de salida del elemento 6 del subgrupo de validación



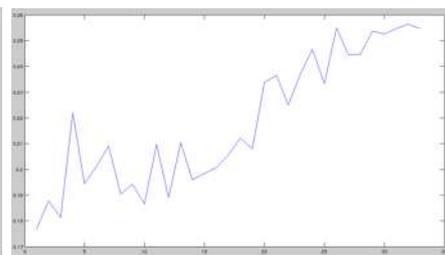
(g) Vector de salida del elemento 7 del subgrupo de validación



(h) Vector de salida del elemento 8 del subgrupo de validación



(i) Vector de salida del elemento 9 del subgrupo de validación



(j) Vector de salida del elemento 10 del subgrupo de validación

Fig. 6.14: Vectores de salida de la red propagando los elementos del subgrupo de validación.





# Capítulo 7

## Conclusiones

La rejilla de difracción utilizada a lo largo del desarrollo experimental tiene resultados favorables al poder distinguir las variaciones en la concentración del analito, sin embargo, aún se puede mejorar en ese aspecto. La limitante principal sobre la rejilla es la función pupila de la cámara considerando los resultados descritos en el capítulo anterior, el espectro obtenido podría caer hasta dos veces más grande permaneciendo dentro de la función pupila. Por ello una mejoría importante para futuros resultados es la construcción de una rejilla con una frecuencia mayor.

Satisfactoriamente la cámara del dispositivo móvil es suficientemente sensible para detectar cambios de concentración del orden de 2 porciones de analito por porción de agua. Se puede ver de forma gráfica en la figura en la cual se muestran los espectros a distintas concentraciones. Esto incita a seguir con la implementación de la red neuronal para poder asociar los espectros de las muestras con su respectiva concentración.

Al confirmar que la eficiencia de la red neuronal es al menos del 80 %, lejos de dar por concluida la tesis podemos dar un amplio horizonte de futuras aplicaciones con los mismos principios discutidos a lo largo de ésta. Como primer punto se tiene que hacer

un diseño de un instrumento que ayude a poder montar el arreglo del espectrofotómetro a cualquier dispositivo móvil de forma versátil y eficaz. Otro punto muy importante será la construcción de otra red neuronal que sea capaz de reconocer distintas concentraciones de más sustancias, en otras palabras sería como tener una base de datos intrínseca en la construcción de la red, que por sí sola representaría un ahorro de memoria significativo en comparación de tener como tal una base de datos sobre cada sustancia analizada. Sin duda alguna una meta a alcanzar será crear una aplicación en Android donde se carguen los datos de la red neuronal para finalmente tener en el dispositivo móvil el espectrofotómetro completo, desde su hardware hasta su software. Estos serían solo algunos puntos que se pueden alcanzar con los resultados obtenidos.

Las ventajas principales, de forma personal, de haber realizado el proyecto de la forma en que se planteó van desde aprender más acerca de los fenómenos de interferencia y difracción así como el conocimiento intelectual y práctico sobre la manipulación de los aparatos y elementos ópticos del laboratorio y su aplicación, hasta la destreza para programar en el lenguaje de MatLab para crear de forma correcta una red neuronal, pues esto conlleva un conocimiento general entre teoría, práctica y programación.

# Apéndice A

## Algoritmo de la función: actualizarPesos

```
1  function [vector ,error ,wn_i ,wn_o]=actualizarPesos (w_i ,w_o ,I ,  
    target ,epsilon)  
2  
3  [o ,p]=size (I);  
4  [itarget ,jtarjet]=size (target);  
5  p=int64 (max(p,o));  
6  itarget=int64 (max(itarget ,jtarjet));  
7  suma_h=0;  
8  
9  for n=1:p;  
10     for i=1:p;  
11         suma_h=suma_h+w_i(n,i)*I(i);  
12     end  
13     Net_h(n)=suma_h+0.2;  
14     suma_h=0;  
15 end  
16  
17 for n=1:p;  
18     Out_h(n)=1/(1+exp(-Net_h(n)));  
19 end  
20  
21 for n=1:itarget;  
22     for i=1:p;  
23         suma_h=suma_h+w_o(n,i)*Out_h(i);  
24     end  
25     Net_o(n)=suma_h+0.3;  
26     suma_h=0;  
27 end  
28  
29 for n=1:itarget;
```

86 APÉNDICE A. ALGORITMO DE LA FUNCIÓN: ACTUALIZAR PESOS

```

30  Out_o(n)= 1/(1+exp(-Net_o(n)));
31  end
32
33  for n=1:itarget;
34    delta_o(n)=(target(n)-Out_o(n))*Out_o(n)*(1-Out_o(n));
35  end
36
37  for n=1:p;
38    for k=1:itarget;
39      suma_h=suma_h+delta_o(k)*w_o(k,n);
40    end
41    delta_i(n)= suma_h*Out_h(n)*(1-Out_h(n));
42    suma_h=0;
43  end
44
45  for n=1:itarget;
46    for i=1:p;
47      partial_w_o(n,i)=epsilon*delta_o(n)*Out_h(i);
48    end
49  end
50
51  for n=1:p;
52    for i=1:p;
53      partial_w_i(n,i)=epsilon*delta_i(n)*I(i);
54    end
55  end
56
57  Error=0;
58  for i=1:itarget;
59    Error=(1/2)*(target(i)-Out_o(i))^2+Error;
60  end
61
62  error=Error;
63
64  for i=1:itarget;
65    vector(i)=Out_o(i);
66  end
67
68  for n=1:itarget;
69    for i=1:p;
70      wn_o(n,i)=w_o(n,i)+partial_w_o(n,i);
71    end
72  end
73
74  for n=1:p;
75    for i=1:p;
76      wn_i(n,i)=w_i(n,i)+partial_w_i(n,i);
77    end
78  end

```

## Apéndice B

### Algoritmo de la función: propagacionAdelante

```
1 function [salida ,error]=propagacionAdelante(w_i ,w_o,I ,target)
2
3 [o ,p]=size(I);
4 [itarget ,jtarjet]=size(target);
5 p=int64(max(p,o));
6 itarget=int64(max(itarget ,jtarjet));
7
8 suma_h=0;
9
10 for n=1:p;
11     for i=1:p;
12         suma_h=suma_h+w_i(n,i)*I(i);
13     end
14     Net_h(n)=suma_h+0.2;
15     suma_h=0;
16 end
17
18 for n=1:p;
19     Out_h(n)=1/(1+exp(-Net_h(n)));
20 end
21
22 for n=1:itarget;
23     for i=1:p;
24         suma_h=suma_h+w_o(n,i)*Out_h(i);
25     end
26     Net_o(n)=suma_h+0.3;
27     suma_h=0;
28 end
29
30 for n=1:itarget;
```

88 APÉNDICE B. ALGORITMO DE LA FUNCIÓN: PROPAGACION ADELANTE

```
31 Out_o(n)= 1/(1+exp(-Net_o(n)));
32 end
33
34 Error=0;
35
36 for i=1:itarget;
37     Error=(1/2)*(target(i)-Out_o(i))^2+Error;
38 end
39
40 error=Error;
41
42 for i=1:itarget;
43 salida(i)=Out_o(i);
44 end
```

# Apéndice C

## Algoritmo de la función: propagacionAtras

```
1 function [vect ,errorV ,errorForV ,wi,wo]= propagacionAtras(iter ,
   errorMax ,M,V,Mfor ,Vfor)
2
3     % [Mfor, Vfor] subgrupo de validacion
4     % [M, V] subgrupo de entrenamiento
5     [o, p]=size(M) ;
6     [q, r]=size(V) ;
7     [oo, pp]=size(Mfor) ;
8     [qq, rr]=size(Vfor) ;
9     epsilon=0.250;
10
11 for i=1:o;
12     for j=1:p;
13         w_i(i, j)=2*rand-1;
14     end
15 end
16
17 for i=1:q;
18     for j=1:r;
19         w_o(i, j)=2*rand-1;
20     end
21 end
22 j=1;
23 errorV=errorMax*1000;
24
25 while errorV>errorMax;
26
27     if j>iter
28         break
29     end
```

90 APÉNDICE C. ALGORITMO DE LA FUNCIÓN: PROPAGACION ATRAS

```

30     if j>2
31         if errorForV(j-1)<errorMax
32             break
33         end
34     end
35     % Tomar elementos aleatorios
36     [m]=RandSimRepeticao(1,p,p);
37     sumaEr=0.0;
38     h=0;
39     sumaErFor=0.0;
40
41     for i=1:p;
42
43         h=h+1;
44
45         [vect , errorn , wn_i , wn_o]=actualizarPesos ( w_i , w_o ,M(: ,m
46             (1 , i)) ,V(: ,m(1 , i)) , epsilon );
47         [VEC, errorFor]=propagacionAdelante ( wn_i , wn_o , Mfor (: , h) ,
48             Vfor (: , h) );
49
50         w_i=wn_i;
51         w_o=wn_o;
52         sumaErFor=sumaErFor+errorFor;
53         sumaEr=sumaEr+errorn;
54
55     if i==p;
56         h=0;
57         errorV(j)=sumaEr/double(p);
58         sumaEr=0;
59
60     end
61     if h==pp;
62         h=0;
63         errorForV(j)=sumaErFor/double(pp);
64         sumaErFor=0;
65     end
66 end
67 wi=w_i;
68 wo=w_o;

```

# Apéndice D

## Algoritmo de la función: AnálisisRedFinal

```
1 function [vect ,a]= AnalisisRedFinal(M,V, w_i , w_o)
2
3     [o ,p]=size (M) ;
4     [oo ,pp]=size (V) ;
5     pp=int64 (pp) ;
6     p=int64 (p) ;
7 for i =1:p
8     [vect (: , i) ,error]=propagacionAdelante (w_i , w_o ,M(: , i) ,V(: , i))
9     ;
10 end
11 suma=0;
12 n=0;
13 k=0;
14 for j =1:pp
15     for i =2:oo-2
16         n=n+1;
17
18         for l =0:3
19
20             suma=suma+((vect (i-1+l , j) )/4) ;
21
22         end
23
24         flot (j ,n)=suma;
25         suma=0;
26
27     end
28     if j<pp
29         n=0;
```

```

30     end
31 end
32 k=1;
33 for i=1:pp
34     for j=1:n
35         if flot(i,j)==max(flots(i,:))
36             a(k,1)=j-1;
37             if j-1==0
38                 a(k,1)=1;
39             end
40             a(k,2)=j+2;
41             if j+2==n
42                 a(k,2)=3;
43             end
44             k=k+1;
45         end
46     end
47 end
48 flg=0;
49 for i=1:pp
50     for j=1:oo
51         if V(j,i)>0.2
52             a(i,3)=j-2;
53             a(i,4)=j;
54         end
55     end
56 end
57 end
58
59 num=0;
60 acerto=0;
61 fallo=0;
62 for i=1:pp
63
64     if a(i,1) >= a(i,3) & a(i,1) <= a(i,4) & num==0;
65         fprintf('El rango donde la salida de la red por el
66             elemento %d es máximo es en [%d,%d] y el valor
67             deseado está en [%d,%d] por lo tanto la red acerto\n',
68             i,a(i,1),a(i,2),a(i,3),a(i,4));
69         num=1;
70         acerto=acerto+1;
71     end
72
73     if a(i,2) <= a(i,4) & a(i,2) >= a(i,3) & num==0;
74         fprintf('El rango donde la salida de la red por el
75             elemento %d es mínimo es en [%d,%d] y el valor
76             deseado está en [%d,%d] por lo tanto la red acerto\n',
77             i,a(i,1),a(i,2),a(i,3),a(i,4));
78         num=4;

```

```
73     acerto=acerto+1;
74 end
75 if num==0;
76     fprintf( 'El rango donde la salida de la red por el
             elemento %d es máximo es en [%d,%d] y el valor
             deseado esta en [%d,%d] por lo tanto la red fallo\n',
             i ,a(i,1) ,a(i,2) ,a(i,3) ,a(i,4));
77     num=5;
78     fallo=fallo+1;
79 end
80 num=0;
81 end
82 efectivo=(double(acerto)/double(pp)*100);
83 fprintf( 'La red acerto en %d locaciones y fallo en %d por lo
             tanto tiene una efectividad del %f por ciento\n',acerto ,
             fallo ,efectivo);
84
85 end
```



# Referencias

- [1] HOLLADAY, L. L. *Optical Society of America. Review of Scientific Instruments*, 1926, vol. 12, no 4.
- [2] LITTER, M. I.; ARMIENTA, M. A.; FARÍAS, S. S. *Metodologías analíticas para la determinación y especiación de arsénico en aguas y suelos*. IBEROARSEN, CYTED, Buenos Aires, Spanish, 2009.
- [3] MARQUINA, JOSÉ ERNESTO, ET AL. *La metodología de Newton. Ciencias*, 2003, no 070.
- [4] PERKAMPUS, HEINZ-HELMUT. *UV-VIS Spectroscopy and its Applications*. Berlin: Springer-Verlag, 1992.
- [5] KOTLER, PHILIP, ET AL. *Marketing*. Pearson Higher Education AU, 2015.
- [6] PALMER, CHRISTOPHER A.; LOEWEN, ERWIN G. *Diffraction grating handbook*. Springfield, Ohio, USA: Newport Corporation, 2005.
- [7] YÁÑEZ-ARANCIBIA, ALEJANDRO; TWILLEY, ROBERT R.; LARA-DOMÍNGUEZ, ANA LAURA. *Los ecosistemas de manglar frente al cambio climático global*. Madera y Bosques, 1998, vol. 4, no 2, p. 3-19.
- [8] FLORÉEN, PATRIK, ET AL. *Multicast time maximization in energy constrained wireless networks. En Proceedings of the 2003 joint workshop on Foundations of mobile computing*. ACM, 2003. p. 50-58.
- [9] SCHALKOFF, ROBERT J. *Artificial neural networks*. McGraw-Hill Higher Education, 1997.
- [10] SHARMA, KAILASH K. *Optics: principles and applications*. Academic Press, 2006.

- [11] REITZ, JOHN R.; MILFORD, FREDERICK J.; CHRISTY, ROBERT W. *Foundations of electromagnetic theory*. Addison-Wesley Publishing Company, 2008.
- [12] HECHT, EUGENE. *Optics*, 4th. International edition, Addison-Wesley, San Francisco, 2002, vol. 3.
- [13] MICHIE, DONALD; SPIEGELHALTER, DAVID J.; TAYLOR, CHARLES C. *Machine learning, neural and statistical classification*. 1994.
- [14] HILERA GONZÁLEZ, JOSÉ RAMÓN, ET AL. *Redes neuronales artificiales: fundamentos, modelos y aplicaciones*. 2000.