UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS
Y DE LA ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA

El Teorema de Green-Tao

TESINA
QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS

P R E S E N T A :
ALBERTO VARGAS RODRÍGUEZ

Director de la Tesina
Dr. Timothy Gendron Thornton
IMATE CUERNAVACA

Ciudad de México Abril 2016

Índice general

Introducción

En matemáticas en particular en la teoría de números, emergen de manera natural las progresiones aritméticas, es decir conjuntos de la forma $\{a+bn\}_{n=0}^{k}$ donde a, b y k son números naturales, y al número k se le suele llamar el tamaño o longitud de la progresión, y en este caso se les acostumbra escribir como *k-progresión aritmética*.
Correspondiente a las progresiones aritméticas uno de los temas que ha sido de gran interés en los últimos años ha sido estudiar progresiones aritméticas sobre algunos conjuntos. En concreto está el problema acerca de lo que necesita cumplir un subconjunto de los naturales para asegurar que siempre podemos hallar una k-progresión aritmética para cualquier $k \in \mathbb{N}$, y en esta dirección está el *Teorema de Szemerédi* [3], que dice lo siguiente:

Teorema 0.1 (Szemerédi) *Sea un conjunto $A \subseteq \mathbb{N}$. Si para cada $N \in \mathbb{N}$ se denota a $[\![N]\!] = \{1, 2, \ldots, N\}$, y si pasa que*

$$\limsup_{N\to\infty} \frac{|A \cap [\![N]\!]|}{N} > 0, \tag{1}$$

entonces el conjunto A contiene progresiones aritméticas de longitud arbitraria, esto es que para cada $k \in \mathbb{N}$, existe una k-progresión aritmética.

Al límite superior obtenido en (1) se le llama *densidad superior* del conjunto A. Sobre uno de los conjuntos que se puede preguntar si es que posee esta propiedad de tener progresiones aritméticas de longitud arbitraria, es el de los números primos que denotamos como $\mathcal{P}$. Y una forma de averiguar esto sería empleando el Teorema de Szemerédi, para esto bastaría con calcular la densidad superior de $\mathcal{P}$. Si $\pi(N) = |\mathcal{P} \cap [\![N]\!]|$, es la cantidad de primos desde 1 hasta N, de acuerdo al Teorema del Número Primo[1] existe una sucesión de números reales $\{C_N\}_{N=1}^{\infty}$ que converge a 0 cuando $N \to \infty$, tal que $\pi(N) = {}^{(1+C_N)N}/_{\ln N}$ para cada $N \in \mathbb{N}$, y por lo tanto

$$\limsup_{N\to\infty} \frac{|\mathcal{P} \cap [\![N]\!]|}{N} = \limsup_{N\to\infty} \frac{\pi(N)}{N} = \limsup_{N\to\infty} \frac{1+C_N}{\ln N} = 0,$$

así que no es posible concluir con el Teorema de Szemerédi, que el conjunto de los números primos contiene progresiones aritméticas de longitud arbitraria. Por muchos años permaneció esta pregunta abierta hasta que Ben Green y Terence Tao en 2006 [4], dieron una respuesta afirmativa, que es el resultado principal del que se hablará en este trabajo:

Teorema 0.2 (Green-Tao [2006]) *El conjunto de los números primos contiene progresiones aritméticas de longitud arbitraria.*

Objetivos de la Tesina

Uno de los objetivos es exponer a lo que actualmente se conoce como el Teorema de Green y Tao, que como ya se mencionó habla acerca de la arbitrariedad en el tamaño de las progresiones aritméticas que podemos hallar en el conjunto de los números primos. Así como comentar los resultados auxiliares que hacen posible su demostración, resultados que en algunos casos

[1] Véase la sección 1.8 de [15]

omitiremos su demostración que pueden llegar a ser algo técnicas, esto con motivo de sobrepasar los límites de esta tesina.
Otro de los objetivos que se planteó fué el recopilar los resultados de la demostración al Teorema de Green-Tao, que proponen David Conlon, Jacob Fox y Yufei Zhao en [1] y [2]. La cual a nuestro punto de vista ofrece numerosas "simplificaciones" a la prueba original de Ben Green y Terence Tao [4]. Simplificaciones no sólo en argumentos sino además en disciplinas y/o enfoques distintos, con las que se tratan algunos de los resultados (por ejemplo hacen uso de algunas nociones de Teoría de Hipergráficas). Se puede constatar que nos ofrece una mejor comprensión y exposición del resultado objetivo, de nuevo teniendo en cuenta que a pesar de estas simplificaciones en algunos casos sólo es posible ofrecer algunos comentarios de sus demostraciones, pues son en la mayoría de los casos técnicas y/o extensas. Por esta razón se han seleccionado presentar las pruebas de aquellos resultados que a nuestro juicio, nos ayudan a entender mejor el teorema.

Preliminares

Al principio de esta sección se mencionó el Teorema de Szemerédi y se hizo notar que no fué posible aplicar este resultado para el conjunto de los números primos, pues su densidad superior al calcularla daba 0. Sin embargo Green y Tao extienden este resultado a subconjuntos de números naturales, que les llaman *pseudoaleatorios* o que cumplen ciertas hipótesis de aleatoriedad. Donde ahora estos conjuntos pseudoaleatorios sean nuestro conjunto base, es decir que tomen el papel de los números naturales en el Teorema de Szemerédi y que ahora los subconjuntos densos, es decir con densidad superior positiva, su densidad se calcule o sea relativa, a estos conjuntos pseudoaleatorios. Y si llegan a cumplir con esta condición de densidad, entonces se puede asegurar que estos "nuevos" conjuntos densos poseen la propiedad de tener progresiones aritméticas de longitud arbitraria.
A esta propiedad que heredan los naturales hacia los conjuntos nombrados pseudoaleatorios, Green y Tao le llaman *principio de transferencia*. De igual forma a la extensión del Teorema de Szemerédi le llamarón *Teorema de Szemerédi Relativo*, que puede enunciarse informalmente asi: «*Si S es un subconjunto de los números naturales que satisface ciertas condiciones de pseudoaleatoriedad y A es un subconjunto de S con densidad relativa positiva, entonces A contiene progresiones aritméticas de longitud arbitraria*». Este teorema es uno de los dos resultados clave para demostrar el Teorema de Green-Tao, mientras que el segundo más que un resultado, es el encontrar un subconjunto de naturales que Green y Tao lo llaman el conjunto de los *casi primos*, donde el conjunto de los números primos (o al menos un subconjunto) tenga densidad relativa positiva y así concluir que contienen progresiones aritméticas de longitud arbitraria.
Para probar el Teorema de Szemerédi Relativo, se necesitan básicamente dos resultados que son el *Teorema del Modelo Denso* y el *Lema de Conteo*.
El Teorema del Modelo Denso hablando coloquialmente, nos permite decir que si $S \subseteq \mathbb{N}$ cumple ciertas condiciones de pseudoaleatoriedad entonces a cualquier subconjunto A de S relativamente denso, se le puede "asociar" un subconjunto $\widetilde{A} \subseteq \mathbb{N}$, denso (en los naturales). Por otro lado el Lema de Conteo relaciona bajo una constante de proporcionalidad, el número de k-progresiones aritméticas en A, con las que hay en $\widetilde{A}$, y como $\widetilde{A}$, tiene densidad superior positiva en $\mathbb{N}$, el Teorema de Szemerédi (Teorema 0.1) implica que $\widetilde{A}$ contiene progresiones aritméticas de longitud arbitraria, lo cual a su vez implicará que el conjunto A también tiene

esta propiedad.
Ahora hablemos un poco de como está estructurado este trabajo. En la primera sección se presenta el Teorema del Modelo Denso y se comenta acerca de su demostración, después se establecen algunos conceptos de la teoría de hipergráficas y aquí se establece la primera definición que captura el concepto de pseudoaleatoriedad que se conocerá como *condición de formas k-lineales*. Y después escribiremos el Lema de Conteo para hipergráficas y al igual haremos algunos comentarios acerca de su demostración; y se concluye la primera sección con la prueba del Teorema de Szemerédi Relativo.
En la segunda sección se explica la "construcción" del conjunto pseudoaleatorio que contiene a los números primos[2], tal que estos sean densos aquí y poder cumplir con las hipótesis del Teorema Relativo de Szemerédi, concluyendo con la demostración del Teorema de Green-Tao.

Ahora vamos a intoducir alguna notación que se emplea en el resto del trabajo. Ya se ha venido empleando el símbolo $\mathbb{N}$ para los números naturales $\{1, 2, 3, \ldots\}$, de igual manera se usarán a los símbolos $\mathbb{Z}$ y $\mathbb{R}$, para denotar a los enteros y reales respectivamente, como usualmente se acostumbra.
También acostumbraremos a denotar $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$, con $N \in \mathbb{N}$, es decir $\mathbb{Z}_N$ será el grupo de clases residuales de los enteros módulo N, en general se considerará a N como número primo, esto con objetivo de que $\mathbb{Z}_N$ sea siempre un campo para poder dividir.
Se usará ampliamente la notación O-grande y o-pequeña, en particular nosotros la emplearemos con funciones cuyo dominio son los naturales, de la siguiente forma: sean $f, g : \mathbb{N} \to \mathbb{R}$ funciones,

i. Decimos que f es O-grande de g cuando n tiende a ∞ y escribimos $f(n) = O(g(n))$ cuando $n \to \infty$, si existen $M > 0$, y $n_0 \in \mathbb{N}$, tal que si $n \geqslant n_0$, entonces $|f(n)| \leqslant M|g(n)|$.

ii. Se dice f es o-pequeña de g cuando n tiende a ∞, y se escribe $f(n) = o(g(n))$ cuando $n \to \infty$, si $\lim_{n\to\infty} |f(n)/g(n)| = 0$.

En ambos caso vamos a suprimir la notación $n \to \infty$, pues al tratar con funciones con dominio en los naturales se sobreentiende que los límites se toman en infinito. En la notación O-grande en algunas ocasiones colocaremos subíndices para denotar alguna dependencia sobre la constante M, por ejemplo $f(n) = O_k(g(n))$, significa que M dependerá del parámetro k. En el caso de la notación o-pequeña los subíndices que lleguen a escribirse, significan en este caso una dependencia de la función f, por ejemplo $f(n) = o_k(g(n))$, indica que la función f también depende del parámetro k.
Una de las herramientas que se usa constantemente en prácticamente todo el trabajo es *la media aritmética* o *esperanza* de una función sobre un conjunto, pues nos permitirá "contar" las progresiones aritméticas que hay en un conjunto dado. Ahora en concreto sea A un conjunto finito, no vacío y $f : A \to \mathbb{R}$ es una función, se define la media aritmética o esperanza de f en A, como

$$\mathbb{E}_{x\in A}[f(x)] = \frac{1}{|A|}\sum_{x\in A} f(x).$$

Nótese que aún si f depende de dos (o más) variables, se puede aplicar esta definición. Por ejemplo si $A, B \neq \varnothing$, son conjuntos finitos y $f : A \times B \to \mathbb{R}$ es función, entonces

[2]De hecho será a un subconjunto de estos, pues no será posible contener al conjunto de los números primos, esto tiene que ver con el hecho de que los primos no son tan aleatorios como en este caso nos gustaría que fueran.

$$\mathbb{E}_{x\in A.y\in B}[f(x,y)] = \frac{1}{|A\times B|}\sum_{(x,y)\in A\times B} f(x,y)$$
$$= \frac{1}{|A||B|}\sum_{x\in A}\sum_{y\in B} f(x,y) = \frac{1}{|A||B|}\sum_{y\in B}\sum_{x\in A} f(x,y)$$
$$= \mathbb{E}_{x\in A}[\mathbb{E}_{y\in B}[f(x,y)]] = \mathbb{E}_{y\in B}[\mathbb{E}_{x\in A}[f(x,y)]],$$

donde sin ningún problema podemos intercambiar la suma, pues tenemos una cantidad finita, de sumandos finitos. También en el caso de que los conjuntos A y B sean iguales se escribirá símplemente la esperanza de f como $\mathbb{E}_{x,y\in A}[f(x,y)](=\mathbb{E}_{x,y\in B}[f(x,y)])$.
Ahora regresando al problema de querer contar progresiones aritméticas o en principio saber si es que hay alguna, supongamos que tenemos un conjunto $A\subseteq\mathbb{N}$, y se considera la *función característica* 1_A del conjunto, donde $1_A(x)=1$, si $x\in A$, y $1_A(x)=0$, si $x\notin A$, a partir de esto tenemos que

$$1_A(x)1_A(x+d)\cdots 1_A(x+(k-1)d) = \begin{cases} 1 & \text{si } x, x+d,\ldots,x+(k-1)d\in A \\ 0 & \text{de otra manera}\end{cases}$$

con $k,d\in\mathbb{N}$. Abusando de la notación se escribirá $A\cap\mathbb{Z}_N$, para referirnos al conjunto $A\cap[\![N]\!]$, reducido módulo N, es decir que a cada elemento del conjunto $A\cap[\![N]\!]$, se le asocia su clase en $\mathbb{Z}_N$, ahora si queremos obtener la cantidad total de k-progresiones aritméticas en $A\cap\mathbb{Z}_N$, se puede calcular la suma

$$\sum_{x,\,d\in\mathbb{Z}_N} 1_A(x)1_A(x+d)\cdots 1_A(x+(k-1)d), \tag{2}$$

notemos que aquí se ha cambiado el dominio de los parámetros, del conjunto de los naturales $\mathbb{N}$ a $\mathbb{Z}_N$, esto principalmete por razones técnicas pues en principio en $\mathbb{Z}_N$ tenemos una estructura de campo y es finito. Aúnque esto tiene algunos ligeros inconvenientes, pues el considerar progresiones aritméticas en $\mathbb{Z}_N$, podría contar progresiones aritméticas de más, como por ejemplo $N-1,0,1$ es una 3-progresión aritmética en $\mathbb{Z}_N$, pero que no lo es en $A\cap[\![N]\!]$. Para evitar esta complicación lo que se hace es considerar a $A\cap[\![N]\!]$, sobre un grupo de residual más "grande" y prevenir que se cuenten este tipo de progresiones aritméticas alrededor del 0.
Por otro lado, observemos que la suma en (2) puede ser vista como una expresión de la siguiente manera $|N|^2\mathbb{E}_{x,d\in\mathbb{Z}_N}[1_A(x)1_A(x+d)\cdots 1_A(x+(k-1)d)]$, luego la suma en (2) es positiva si y sólo si

$$\mathbb{E}_{x,d\in\mathbb{Z}_N}[1_A(x)1_A(x+d)\cdots 1_A(x+(k-1)d)] > 0, \tag{3}$$

por lo tanto la manera de averiguar si existe al menos una progresión aritmética en un conjunto dado, es investigar cuándo o bajo que condiciones es positiva expresiones como estas.
Regresemos al Teorema de Szemerédi, para ver como se puede relacionar con lo visto anteriormente, y primero escribiremos la versión original de Szemerédi [3], que es equivalente a la del Teorema 0.1, la cual es.

Teorema 0.3 *Sean $\delta>0$ y $k\geq 3$ un entero, donde ambos son parámetros fijos. Entonces existe un entero minimal $N_0(\delta,k)<\infty$, con la siguiente propiedad. Si $N\geq N_0(\delta,k)$ y $A\subseteq\mathbb{Z}_N$ es cualquier conjunto tal que $|A|\geq\delta N$, entonces A contiene una k-progresión aritmética.*

Aunque su prueba original emplea técnicas de combinatoria, actualmente se pueden hallar diversas pruebas en distintas áreas, una de ellas emplea teoría ergódica y se le debe a Furstenberg [6]. Tal y como se menciona en el artículo de Green y Tao, la teoría ergódica sugiere la siguiente forma del Teorema de Szemerédi.

Teorema 0.4 *Para cada $N \in \mathbb{N}$, se denota $\nu_{const}^{(N)} : \mathbb{Z}_N \to \mathbb{R}$, la función constante $\nu_{const}^{(N)} \equiv 1$. Sean $0 < \delta \leqslant 1$, $k \in \mathbb{N}$ fijos. Si $f^{(N)} : \mathbb{Z}_N \to [0, \infty)$, es una función tal que $0 \leqslant f^{(N)}(x) \leqslant \nu_{const}^{(N)}(x)$, para cada $x \in \mathbb{Z}_N$, y cumple que $\mathbb{E}_{x \in \mathbb{Z}_N}[f^{(N)}(x)] \geqslant \delta$, entonces*

$$\mathbb{E}_{x,d \in \mathbb{Z}_N}[f^{(N)}(x) f^{(N)}(x+d) \cdots f^{(N)}(x+(k-1)d)] \geqslant c(k,\delta) - o_{k,\delta}(1) \tag{4}$$

donde $c(k,\delta) > 0$, es una constante que no depende de $f^{(N)}$ o N, y $o_{k,\delta}(1)$ denota una función que depende de N y converge a cero, conforme N tiende a infinito, y la elección de esta función puede depender de los parámetros k y δ.

Con métodos de combinatoria puede deducirse este teorema del Teorema 0.3, llevado a cabo por Varnavides [7], aúnque una prueba directa debido a Terence Tao se puede encontrar en [8]. Y es esta la versión que se emplea para demostrar al Teorema de Szemerédi Relativo.
Se decidió escribir estás dos formas equivalentes del Teorema de Szemerédi para hacer notar que en el caso del Teorema 0.3, se asemeja bastante al Teorema 0.1, que se escribió al inicio de la sección, por su lado el Teorema 0.4, deja de lado los conjuntos y trata ahora con funciones, que en cierta forma es comprensible pues es más sencillo trabajar con funciones que con conjuntos. De hecho anteriormente ya se había pasado de un conjunto a una función, y era que le asociabamos su función característica. Más aún la esperanza en (4) se asemeja a la esperanza tomada en (3), así que la ecuación (4) se puede decir que generaliza la forma de contar a las progresiones aritméticas y el Teorema de Szemerédi Relativo es justamente generalizar al Teorema 0.4, donde ya no sólo se considere a funciones f acotadas (o como menciona el artículo de Conlon, Fox y Zhao "densas") sino que incluso ya no necesariamente tengan que ser acotadas (o "dispersas"), esto a su vez la función ν que mayora ya no será tan simple como la función ν_{const}, y se le deberá imponer condiciones para que siga siendo válida la conclusión.

1 | Teorema de Szemerédi Relativo

En esta sección se presentan al Teorema del Modelo Denso y el Lema de Conteo que esencialmente son los dos resultados que se requieren para la demostración del Teorema de Szemerédi Relativo que se expondrá.

Teorema del Modelo Denso y el Lema de Conteo

En la sección anterior en los preliminares, se había hablado en general acerca de lo que el Teorema del Modelo Denso iba a significar para nuestros propositos. Sin tratar de ser demasiado precisos, se pretende explicar más en concreto como es que nos será de utilidad este teorema para demostrar al Teorema Relativo de Szemerédi. Para esto supongamos que tenemos $A \subseteq S \subseteq \mathbb{N}$ subconjuntos. Recordemos que el problema a resolver es saber si hay alguna k-progresión aritmética en el conjunto A. Para saber si la hay o no, se cosideran cada uno de los conjuntos finitos $A \cap \mathbb{Z}_N$, y de una manera adecuada a estos conjuntos se le asocia una función $f_A : \mathbb{Z}_N \to [0, \infty)$, decimos de manera adecuada en el sentido de que pondera a los elementos del conjunto de forma que cumplan ciertas hipótesis de densidad respecto a S. Y de igual manera a los conjuntos $S \cap \mathbb{Z}_N$, se le asocia una función $\nu_S : \mathbb{Z}_N \to [0, \infty)$. Ahora el principal problema es que las funciones f_A pueden no ser acotadas, pero si se puede esperar que como S contiene a A, que ν_S mayore a f_A, es decir que $f_A \leqslant \nu_S$, y bajo estas condiciones puede actuar el Teorema del Modelo Denso de manera que a la función f_A, se "modela" con una función $\widetilde{f_A} : \mathbb{Z}_N \to [0, 1]$, donde notemos que $\widetilde{f_A}$ ahora es una función acotada y es suceptible de aplicarle el Teorema de Szemerédi (Teorema 0.4). Por modelar a f_A con $\widetilde{f_A}$ nos referimos a que $\mathbb{E}_{x \in \mathbb{Z}_N}[\widetilde{f_A}(x)] = \mathbb{E}_{x \in \mathbb{Z}_N}[f_A(x)]$, y que bajo cierta norma $\|\cdot\|$, que se le llama "norma de corte", se tenga que la diferencia (distancia) $\|f_A - \widetilde{f_A}\|$ puede hacerse tan pequeña como se desee, esto con fines de que f_A y $\widetilde{f_A}$ tengan o cuenten una cantidad similar de progresiones aritméticas en A.

Empezemos fijando alguna notación que vamos a usar. Sea $r \in \mathbb{N}$, si $\{X_i\}_{i=1}^r$, es una colección finita de conjuntos, se denotará a $\mathbf{X}_{-i}$ por el conjunto $\mathbf{X}_{-i} = X_1 \times \cdots \times X_{i-1} \times X_{i+1} \times \cdots \times X_r$, para cada $i \in [\![r]\!]$. Y de manera semejante se denotará a $\mathbf{x}_{-i}$ como la siguiente $r-1$-tupla ordenada $\mathbf{x}_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_r) \in \mathbf{X}_{-i}$, esto para cada $i \in [\![r]\!]$. Ahora se define a la *norma de corte* para funciones definidas en $\mathbb{Z}_N$.

Definición 1.1 (Norma de Corte para funciones con dominio en $\mathbb{Z}_N$)
Sean $N, r \in \mathbb{N}$ y $f : \mathbb{Z}_N \to \mathbb{R}$ una función, se define la ***norma de corte*** de f denotada por $\|f\|_{\square,r}$, como el número real

$$\|f\|_{\square,r} = \sup \left| \mathbb{E}_{\mathbf{x} \in \mathbb{Z}_N^r}[f(\phi(\mathbf{x}))1_{A_1}(\mathbf{x}_{-1})1_{A_2}(\mathbf{x}_{-2}) \cdots 1_{A_r}(\mathbf{x}_{-r})] \right|,$$

donde $\phi : \mathbb{Z}_N^r \to \mathbb{R}$, está definida por $\phi(\mathbf{x}) = \phi(x_1, \ldots, x_r) = x_1 + \cdots + x_r$, y el supremo se toma sobre todos los conjuntos $A_1, \ldots, A_r \subseteq \mathbb{Z}_N^{r-1}$.

No es complicado comprobar que en efecto la norma de corte, es en realidad una norma sobre el espacio de funciones $\mathbb{R}^N = \{f : \mathbb{Z}_N \to \mathbb{R} |\ f \text{ es función }\}$. A continuación se escribe el *Teorema del Modelo Denso*.

Teorema 1.1 (del Modelo Denso) *Sean $N, r \in \mathbb{N}$. Para todo $\varepsilon > 0$, existe $\varepsilon' > 0$, tal que si $\nu : \mathbb{Z}_N \to [0, \infty)$, con $\|\nu - 1\|_{\square,r} \leqslant \varepsilon'$, entonces para cada función $f : \mathbb{Z}_N \to [0, \infty)$, con $\mathbb{E}_{x\in\mathbb{Z}_N}[f(x)] \leqslant 1$, y $f \leqslant \nu$, existe una función $\widetilde{f} : \mathbb{Z}_N \to [0, 1]$, que cumple con*

i. $\mathbb{E}_{x\in\mathbb{Z}_N}[\widetilde{f}(x)] = \mathbb{E}_{x\in\mathbb{Z}_N}[f(x)]$, y

ii. $\|f - \widetilde{f}\|_{\square,r} \leqslant \varepsilon$. □

Para la prueba de este teorema se emplea una versión más general del mismo, debido a Green, Tao y Ziegler (el enunciado de este teorema, con su demostración, propuesta por Reingold, Trevisan, Tulsiani y Vadhan, se puede encontrar en [12]). Hay que usar el hecho de que la norma de corte de una función f se puede expresar de la siguiente forma $\|f\|_{\square,r} = \sup\left\{ \left|\mathbb{E}_{x\in\mathbb{Z}_N}[f(x)\varphi(x)]\right| \,\middle|\, \varphi \in \mathcal{F}_r \right\}$, donde $\mathcal{F}_r \subseteq \mathbb{R}^N$, es la familia de funciones que pueden ser escritas como combinación convexa, de la convolución de r funciones caractrísticas, de conjuntos en $\mathbb{Z}_N^{r-1}$, probando que $\mathcal{F}_r$ es cerrado bajo la multiplicación de funciones, y despues se usa la versión del Teorema del Modelo Denso de Green, Tao y Ziegler.
Teniendo a la mano el Teorema del Modelo Denso, se presentan a partir de aquí algunos conceptos que son el de hipergráfica e hipergráfica cargada. Advirtiendo que la manera en que se presentan aquí, no es la forma estandar que se presenta usualmente en la teoría de hipergráficas, pues se ha adapatado la notación a nuestra conveniencia para nuestros propositos, y es en esta parte que la demostración expondremos de Conlon, Fox y Zhao, que se diferencía de la otorgada por Ben Green y Terence Tao, y se hará más notorio en la condición de formas lineales (Definición 1.5 y 1.6) y en el Lema de Conteo 1.2.
Primero introduzcamos más notación y fijemos $k, r \in \mathbb{N}$, con $r \leqslant k$. Ahora sea $\{X_i\}_{i=1}^k$ una colección de k conjuntos (cualesquiera) finitos y no vacios. Denotamos al conjunto de índices $\mathcal{I}_r$, como

$$\mathcal{I}_r = \left\{(i_1, i_2, \ldots, i_r) \in [\![k]\!]^r \,\middle|\, 1 \leqslant i_j < i_{j+1} \leqslant k,\ \forall j \in [\![r-1]\!]\right\},$$

y si $\alpha \in \mathcal{I}_r$ es de la forma $\alpha = (i_1, i_2, \ldots, i_r)$, se denotará como $\bar{\boldsymbol{X}}_\alpha$ o $\bar{\boldsymbol{X}}_{(i_1,i_2,\ldots,i_r)}$ al producto cartesiano $\bar{\boldsymbol{X}}_\alpha = X_{i_1} \times X_{i_2} \times \cdots \times X_{i_r}$. También se define $\boldsymbol{X}_\alpha = \boldsymbol{X}_{(i_1,i_2,\ldots,i_r)} = \{\alpha\} \times \bar{\boldsymbol{X}}_\alpha$, y la unión de todos estos conjuntos para r fija, se denotará por $\bigcup_r[\{X_i\}_{i=1}^k]$, es decir

$$\bigcup_r[\{X_i\}_{i=1}^k] = \bigcup_{\alpha\in\mathcal{I}_r} \boldsymbol{X}_\alpha.$$

En la práctica, principalmente consideraremos el caso en que $k \geqslant 3$, $r = k - 1$ y $X_i = \mathbb{Z}_N$ para todo $i \in [\![k]\!]$, en consecuencia $\boldsymbol{X}_\alpha = \{\alpha\} \times \mathbb{Z}_N^{k-1}$ para cada $\alpha \in \mathcal{I}_{k-1}$, y su unión es $\bigcup_{k-1}[\{X_i\}_{i=1}^k] = \mathcal{I}_{k-1} \times \mathbb{Z}_N^{k-1}$.
A partir de esta notación, podemos definir lo que es una *hipergráfica k-partita, r-regular*.

Definición 1.2 (Hipergráfica *k*-partita, *r*-regular)
Sean $k, r \in \mathbb{N}$, con $r \leqslant k$. Si $V = \{X_i\}_{i=1}^k$ una colección finita de conjuntos finitos, no vacios, entonces decimos que la terna $\mathcal{H} = (V, r, E)$, donde $E \subseteq \bigcup_r[\{X_i\}_{i=1}^k]$, es una ***hipergráfica k-partita, r-regular***. A la colección V, se le suele llamar los ***vértices***[3] de la hipergráfica, mientras que al conjunto E se le llaman las ***aristas*** de $\mathcal{H}$.

[3] Estrictamente hablando un vértice es un elemento de X_i, mientras que X_i es un conjunto de vértices, y V es una colección de (conjuntos de) vértices. Pero seguiremos abusando del lenguaje y nos referimos a V como los vértices de la hipergráfica.

El parámetro k, por supuesto indica la cantidad de conjuntos en los que están distribuidos los vértices, donde en cada uno de estos conjuntos no hay dos vértices contenidos en una misma arista. Mientras que el parámetro r indicará la cantidad de vértices que contiene cualquier arista. Y en el caso particular de que una hipergráfica sea 2-regular, coincide con la definición de una gráfica k-partita, pues en este caso las aristas consisten de dos vértices.
Ahora se dará paso a definir lo que entenderemos por una *hipergráfica cargada* (o *ponderada*).

Definición 1.3 (Hipergráfica Cargada)
A la pareja $(\mathcal{H}, g)$ donde $\mathcal{H} = (V, r, E)$ es una hipergráfica k-partita, r-regular con vértices $V = \{X_i\}_{i=1}^k$ y $g : \bigcup_r[\{X_i\}_{i=1}^k] \to [0, \infty)$ es una función en r variables, donde si $\bigcup_r[\{X_i\}_{i=1}^k]\backslash E \neq \varnothing$ entonces $g(x) = 0$, para cada $x \in \bigcup_r[\{X_i\}_{i=1}^k]\backslash E$, se le llamará ***hipergráfica cargada***, y a la función g, se le llama ***función de carga***.

Usualmente dada una colección de funciones $\{g_\alpha\}_{\alpha\in\mathcal{I}_r}$, donde $g_\alpha : \bar{\boldsymbol{X}}_\alpha \to [0, \infty)$ para cada $\alpha \in \mathcal{I}_r$, induciremos una función de carga $g : \bigcup_r[\{X_i\}_{i=1}^k] \to [0, \infty)$, de manera que si $(\alpha, \mathbf{x}) \in \boldsymbol{X}_\alpha$ se define $g(\alpha, \mathbf{x}) = g_\alpha(\mathbf{x})$. Se abusará de la notación escribiendo $g = \{g_\alpha\}_{\alpha\in\mathcal{I}_r}$, y se deben pensar a las funciones g_α como la reestricción de g al conjunto $\boldsymbol{X}_\alpha$, es decir (abusando de la notación) que $g_\alpha = g|_{\boldsymbol{X}_\alpha}$. Y si $\alpha = (i_1, i_2, \ldots, i_r)$, también se acostumbra a g_α escribir como $g_{X_{i_1}\cdots X_{i_r}}$.
Al igual que se hizo para las funciones definidas sobre $\mathbb{Z}_N$, se puede definir una norma de corte para hipergráficas cargadas, o más bien para las funciones de carga, de la siguiente manera.

Definición 1.4 (Norma de Corte para funciones de carga)
Sea $(\mathcal{H}, g)$, una hipergráfica cargada, con $\mathcal{H} = (\{X_i\}_{i=1}^k, r, E)$, y $g = \{g_\alpha\}_{\alpha\in\mathcal{I}_r}$.
Si para cada $\alpha = (i_1, i_2, \ldots, i_r) \in \mathcal{I}_r$, y $g_{X_{i_1}\cdots X_{i_r}} : \boldsymbol{X}_\alpha \to [0, \infty)$, se define

$$\|g_{X_{i_1}\cdots X_{i_r}}\|_\square = \sup\left|\mathbb{E}_{\mathbf{x}\in\boldsymbol{X}_\alpha}\left[g(\mathbf{x})1_{A_1}(\mathbf{x}_{-1})1_{A_2}(\mathbf{x}_{-2})\cdots 1_{A_r}(\mathbf{x}_{-r})\right]\right|, \tag{5}$$

donde el supremo se toma sobre todos los conjuntos posibles $A_j \subseteq \mathbf{X}_{-i_j}$, donde recordemos que $\mathbf{X}_{-i_j} = X_{i_1} \times \cdots \times X_{i_{j-1}} \times X_{i_{j+1}} \times \cdots \times X_{i_r}$, entonces a la ***norma de corte*** para la función de carga $\boldsymbol{g}$ denotada por $\|g\|_\square$, se define como el número real

$$\|g\|_\square = \max_{(i_1,i_2,\ldots,i_r)\in\mathcal{I}_r}\{\|g_{X_{i_1}\cdots X_{i_r}}\|_\square\}.$$

Ahora escribiremos la definición de lo que entenderemos por que una hipergráfica (o más bien su función de carga) cumpla con la *condición de formas k-lineales*, que es una forma más débil de la condición de formas lineales adoptada por Green y Tao (Definición 3.1), en [4].

Definición 1.5 (Condición de formas k-lineales para funciones de carga)
Sea $k \in \mathbb{N}$. Para cada $N \in \mathbb{N}$, sea $\mathcal{H}^{(N)} = (\{X_i^{(N)}\}_{i=1}^k, k-1, E^{(N)})$, una hipergráfica k-partita, $k-1$-regular, con función de carga $\nu^{(N)} : \bigcup_{k-1}[\{X_i^{(N)}\}_{i=1}^k] \to [0, \infty)$, entonces diremos que la colección de hipergráficas cargadas $\{(\mathcal{H}^{(N)}, \nu^{(N)})\}_{N\in\mathbb{N}}$, cumple con la ***condición de formas k-lineales***, si

$$\mathbb{E}_{x_1^{(0)},x_1^{(1)}\in X_1^{(N)},\ldots,x_k^{(0)},x_k^{(1)}\in X_k^{(N)}}\Big[\prod_{i=1}^k\prod_{\omega\in\{0,1\}^k}\nu^{(N)}(\mathbf{x}_{-i}^{(\omega_{-i})})\Big] = 1 + o_k(1), \tag{6}$$

donde $\omega = (\omega_1, \ldots, \omega_k)$, con $\omega_i \in \{0,1\}$, para cada $i \in [\![k]\!]$, y en este caso $\mathbf{x}_{-i}^{(\omega_{-i})} = (x_1^{(\omega_1)}, \ldots, x_{i-1}^{(\omega_{i-1})}, x_{i+1}^{(\omega_{i+1})}, \ldots, x_k^{(\omega_k)})$, donde $x_i^{(\omega_i)} \in X_i^{(N)}$, para cada $i \in [\![k]\!]$. Y que además la igualdad (6) se siga cumpliendo aún cuando sea eliminado cualquier subconjunto de factores $\nu^{(N)}$, de los $k2^{k-1}$ que hay.
Recordando que en (6), como en los demás enunciados siguientes, se escribirá $o_k(1)$, para indicar una función que tiende a cero, cuando $N \to \infty$, y que la elección de esta función puede depender de k.

Con esta terminología podemos escribir el Lema de Conteo, estudiado por Conlon, Fox y Zhao ([1],[2]).

Teorema 1.2 (Lema de Conteo)
Sea $k \in \mathbb{N}$, con $k \geqslant 3$. Para cada $N \in \mathbb{N}$, sea $\mathcal{H}^{(N)} = \{V^{(N)}, k-1, E^{(N)}\}$, una hipergráfica k-partita, $k-1$-regular, con vértices $V^{(N)} = \{X_i^{(N)}\}_{i=1}^k$, y funciones de carga $\nu^{(N)}, g^{(N)}, \widetilde{g}^{(N)} : \bigcup_{k-1}[\{X_i^{(N)}\}_{i=1}^k] \to [0,\infty)$, tal que $\{\nu^{(N)}\}_{N\in\mathbb{N}}$, satisface la condición de formas k-lineales, $g^{(N)} \leqslant \nu^{(N)}$, y $\widetilde{g}^{(N)} \leqslant 1$. Si $\|g^{(N)} - \widetilde{g}^{(N)}\|_\square = o(1)$, entonces

$$\left| \mathbb{E}_{\mathbf{x}\in X_1^{(N)}\times\cdots\times X_k^{(N)}} \Big[\prod_{i=1}^k g^{(N)}(\mathbf{x}_{-i}) - \prod_{i=1}^k \widetilde{g}^{(N)}(\mathbf{x}_{-i}) \Big] \right| = o(1). \tag{7}$$

Su prueba se puede encontrar en [1] y en [2], aunque no es una demostración sencilla, ni corta, pues se necesitan emplear al menos otros dos resultados auxiliares, uno de ellos es una versión más débil de este, donde básicamente se pide que $\nu^{(N)} \equiv 1$ (y en este caso trivialmente $\{\nu^{(N)}\}_{N\in\mathbb{N}}$ satisface la condición de formas k-lineales), es decir donde cada función de carga $g^{(N)}$ también es acotada por 1.

Prueba del Teorema de Szemerédi Relativo

En esta segunda parte de la sección se presentará el Teorema de Szemerédi Relativo, empleando los resultados de la parte previa. Sin embargo aún resta por considerar una definición análoga a la condición de formas lineales para cargas de hipergráficas, pero esta vez para funciones con dominio en $\mathbb{Z}_N$, y que de igual forma será una condición más débil que a las consideradas por Green y Tao en [4].

Definición 1.6 (Condición de Formas Lineales para funciones con dominio en $\mathbb{Z}_N$)
Sea $k \in \mathbb{N}$, con $k \geqslant 3$. Para cada $N \in \mathbb{N}$, sea $\nu^{(N)} : \mathbb{Z}_N \to [0,\infty)$, una función. Se dice que la colección de funciones $\{\nu^{(N)}\}_{N\in\mathbb{N}}$, satisface la ***condición de formas k-lineales*** si

$$\mathbb{E}_{x_1^{(0)},x_1^{(1)},\ldots,x_k^{(0)},x_k^{(1)}\in\mathbb{Z}_N} \Big[\prod_{j=1}^k \prod_{\omega\in\{0,1\}^k} \nu^{(N)}\Big(\sum_{i=1}^k (j-i) x_i^{(\omega_i)} \Big)^{n_{j,\omega}} \Big] = 1 + o_k(1), \tag{8}$$

donde $\omega = (\omega_1, \ldots, \omega_k)$, con $\omega_i \in \{0,1\}$, y que además la identidad en (8) se cumpla para cualquier elección de exponentes $n_{j,\omega} \in \{0,1\}$.

La elección de las formas lineales $f_j : \mathbb{Z}_N^k \to \mathbb{R}$ de la forma $f_j(\mathbf{x}) = \sum_{i=1}^k (j-i)x_i$, con $j \in [\![k]\!]$ son propuestas así como se verá más adelante en la prueba del Teorema de Szemerédi Relativo, con el fin de modelar una k-progresión aritmética.
El siguiente lema se emplea para la prueba del Teorema Relativo de Szemerédi.

Lema 1.1 *Sea $k \in \mathbb{N}$, con $k \geqslant 3$, y para cada $N \in \mathbb{N}$, sea $\nu^{(N)} : \mathbb{Z}_N \to [0,\infty)$, una función. Si $\{\nu^{(N)}\}_{N\in\mathbb{N}}$, satisface la condición de formas k-lineales, entonces*

$$\|\nu^{(N)} - 1\|_{\square,k-1} = o_k(1). \qquad \square$$

En la prueba de este resultado se debe usar $k-1$ veces la desigualdad de Cauchy-Schwarz, y en un momento emplear un cambio de variable justificado por el hecho de que $\mathbb{Z}_N$ es un campo, además de que N es primo relativo[4] a cada elemento en $[\![k]\!]$. Y emplear la condición de formas k-lineales de $\{\nu^{(N)}\}_{n\in\mathbb{N}}$. Así se está en condiciones de presentar una prueba al Teorema Realtivo de Szemerédi.

Teorema 1.3 (Relativo de Szemerédi)
Sea $k \in \mathbb{N}$, con $k \geqslant 3$, si para cada $N \in \mathbb{N}$, tenemos $\nu^{(N)} : \mathbb{Z}_N \to [0,\infty)$ función, tal que la colección de todas ellas $\{\nu^{(N)}\}_{N\in\mathbb{N}}$, satisface la condición de formas k-lineales. Entonces para todo $\delta > 0$, existe $c = c(k,\delta) > 0$, tal que para cada colección de funciones $\{f^{(N)}\}_{N\in\mathbb{N}}$, con $f^{(N)} : \mathbb{Z}_N \to [0,\infty)$, que cumpla $f^{(N)} \leqslant \nu^{(N)}$ y $\mathbb{E}_{x\in\mathbb{Z}_N}[f^{(N)}(x)] \geqslant \delta$ para cada $N \in \mathbb{N}$, se satisface que

$$\mathbb{E}_{x,d\in\mathbb{Z}_N}\left[f^{(N)}(x)\,f^{(N)}(x+d)\cdots f^{(N)}(x+(k-1)d)\right] \geqslant c(k,\delta) - o_{k,\delta}(1). \tag{9}$$

Demostración: Sea $\delta > 0$, que sin pérdida de generalidad podemos suponer que $\delta \leqslant 1$. Ahora como la colección de funciones $\{\nu^{(N)}\}_{N\in\mathbb{N}}$, cumple la condición de formas k-lineales se tiene por el Lema 1.1, que $\|\nu^{(N)} - 1\|_{\square,k-1} = o_k(1)$.
Sea $\{f^{(N)}\}_{N\in\mathbb{N}}$, una sucesión de funciones con $f^{(N)} : \mathbb{Z}_N \to [0,\infty)$, tal que $f^{(N)} \leqslant \nu^{(N)}$ y $\mathbb{E}_{x\in\mathbb{Z}_N}[f^{(N)}(x)] \geqslant \delta$ para cada $N \in \mathbb{N}$, a esta sucesión le asociamos a la sucesión normalizada $\{\widehat{f}^{(N)}\}_{N\in\mathbb{N}}$, donde $\widehat{f}^{(N)} \equiv M_N f^{(N)}$ y $M_N = \delta/\mathbb{E}_{x\in\mathbb{Z}_N}[f^{(N)}(x)]$ para cada $N \in \mathbb{N}$. Observemos que $0 < M_N \leqslant 1$ y $\mathbb{E}_{x\in\mathbb{Z}_N}[\widehat{f}^{(N)}(x)] = \delta \leqslant 1$. Luego por el Teorema del Modelo Denso 1.1, existe una sucesión de funciones $\{\widetilde{f}^{(N)}\}_{N\in\mathbb{N}}$, con $\widetilde{f}^{(N)} : \mathbb{Z}_N \to [0,1]$ para cada $N \in \mathbb{N}$, tal que $\|\widehat{f}^{(N)} - \widetilde{f}^{(N)}\|_{\square,k-1} = o_k(1)$ y $\mathbb{E}_{x\in\mathbb{Z}_N}[\widetilde{f}^{(N)}(x)] = \mathbb{E}_{x\in\mathbb{Z}_N}[\widehat{f}^{(N)}(x)] = \delta$.
Ahora por otro lado, sean $X_1^{(N)} = X_2^{(N)} = \cdots = X_k^{(N)} = \mathbb{Z}_N$, y para cada $j \in [\![k]\!]$ sea $\psi_j^{(N)} : \mathbf{X}_{-j}^{(N)} \to \mathbb{Z}_N$, la función lineal definida por

$$\psi_j^{(N)}(\mathbf{x}_{-j}) = \psi_j^{(N)}(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_k) = \sum_{i\in[k]\setminus\{j\}} (j-i)x_i,$$

para cada $\mathbf{x}_{-j} \in \mathbf{X}_{-j}^{(N)}$.
Además para cada $N \in \mathbb{N}$, sea $\mathcal{H}^{(N)} = (\{X_i^{(N)}\}_{i=1}^k, k-1, E^{(N)})$, una hipergráfica k-partita, $k-1$-regular con $E^{(N)} = \bigcup_{k-1}[\{X_i^{(N)}\}_{i=1}^k]$, y funciones de carga $\lambda^{(N)}, g^{(N)}, \widetilde{g}^{(N)} : \bigcup_r[\{X_i^{(N)}\}_{i=1}^k] \to [0,\infty)$ inducidas[5] respectivamente, a traves de las funciones $\lambda_{-j}^{(N)}, g_{-j}^{(N)}, \widetilde{g}_{-j}^{(N)} : \mathbf{X}_{-j}^{(N)} \to [0,\infty)$ definidas como

$$\lambda_{-j}^{(N)}(\mathbf{x}_{-j}) = \nu^{(N)}(\psi_j^{(N)}(\mathbf{x}_{-j})),$$
$$g_{-j}^{(N)}(\mathbf{x}_{-j}) = \widehat{f}^{(N)}(\psi_j^{(N)}(\mathbf{x}_{-j})),$$

[4] Recordemos que desde el inicio de este trabajo estamos considerando que en general N, siempre es un primo (mucho) mayor que k.

[5] Por ejemplo $g^{(N)} = \{g_{-j}^{(N)}\}_{j=1}^k$, y el subíndice $-j$, representa al elemento $\alpha_j \in \mathcal{I}_{k-1}$ donde $\alpha_j = (1,2,\ldots,j-1,j+1,\ldots,k)$.

$$\widetilde{g}_{-j}^{(N)}(\mathbf{x}_{-j}) = \widetilde{f}^{(N)}(\psi_j^{(N)}(\mathbf{x}_{-j})),$$

para cada $\mathbf{x}_{-j} \in \mathbf{X}_{-j}^{(N)}$, y $j \in [\![k]\!]$. A partir de esto se afirma que

$$\|g_{-j}^{(N)} - \widetilde{g}_{-j}^{(N)}\|_{\square} = \|\widehat{f}^{(N)} - \widetilde{f}^{(N)}\|_{\square,k-1},$$

para cada $j \in [\![k]\!]$, y $N \in \mathbb{N}$. Verifiquemos esta identidad con el caso particular $k = j = 4$, pues el caso general sigue el mismo procedimiento. Tenemos por definición de g_{-4} y $\widetilde{g}_{-4}^{(N)}$ que

$$\begin{aligned}
&\|g_{-4}^{(N)} - \widetilde{g}_{-4}^{(N)}\|_{\square} \\
&= \sup_{A_1,A_2,A_3\subseteq\mathbb{Z}_N^2} \left|\mathbb{E}_{x_1,x_2,x_3\in\mathbb{Z}_N}\left[(\widehat{f}^{(N)}-\widetilde{f}^{(N)})(3x_1+2x_2+x_3)1_{A_1}(x_2,x_3)1_{A_2}(x_1,x_3)1_{A_3}(x_1,x_2)\right]\right|
\end{aligned}$$

empleando los cambios de variable $u_1 = 3x_1$, $u_2 = 2x_2$, y $u_3 = x_3$, quedaría

$$\begin{aligned}
&= \sup_{A_1,A_2,A_3\subseteq\mathbb{Z}_N^2} \left|\mathbb{E}_{u_1,u_2,u_3\in\mathbb{Z}_N}\left[(\widehat{f}^{(N)}-\widetilde{f}^{(N)})(u_1+u_2+u_3)1_{A_1}\left(\tfrac{u_2}{2},u_3\right)1_{A_2}\left(\tfrac{u_1}{3},u_3\right)1_{A_3}\left(\tfrac{u_1}{3},\tfrac{u_2}{2}\right)\right]\right| \\
&= \sup_{A_1,A_2,A_3\subseteq\mathbb{Z}_N^2} \left|\mathbb{E}_{u_1,u_2,u_3\in\mathbb{Z}_N}\left[(\widehat{f}^{(N)}-\widetilde{f}^{(N)})(u_1+u_2+u_3)1_{\frac{A_1}{2}}\left(u_2,u_3\right)1_{\frac{A_2}{3}}\left(u_1,u_3\right)1_{\frac{A_3}{2\cdot 3}}\left(u_1,u_2\right)\right]\right|
\end{aligned}$$

donde $\frac{A_1}{2} = \{(x/2, y) \in \mathbb{Z}_N^2 | (x, y) \in A_1\}$, $\frac{A_2}{3} = \{(x/3, y) \in \mathbb{Z}_N^2 | (x, y) \in A_2\}$ y $\frac{A_3}{2\cdot 3} = \{(x/2, y/3) \in \mathbb{Z}_N^2 | (x, y) \in A_3\}$. Cabe destacar que hacer estos cambios de variable, y la definición de estos conjuntos se pueden hacer siempre y cuando N sea primo relativo a 2 y a 3 (en el caso general se debe pedir que sea primo relativo[6] a $2, 3, \ldots, k-1$), para poder realizar la división por estos números, lo cual se consigue sin ningún problema suponiendo que N, es un primo mayor o igual a k. Por lo tanto

$$\begin{aligned}
&\|g_{-4}^{(N)} - \widetilde{g}_{-4}^{(N)}\|_{\square} \\
&= \sup_{B_1,B_2,B_3\subseteq\mathbb{Z}_N^2} \left|\mathbb{E}_{u_1,u_2,u_3\in\mathbb{Z}_N}\left[(\widehat{f}^{(N)}-\widetilde{f}^{(N)})(u_1+u_2+u_3)1_{B_1}\left(u_2,u_3\right)1_{B_2}\left(u_1,u_3\right)1_{B_3}\left(u_1,u_2\right)\right]\right| \\
&= \|\widehat{f}^{(N)} - \widetilde{f}^{(N)}\|_{\square,3},
\end{aligned}$$

como se quería probar.

Por lo tanto podemos inferir que a partir de este caso particular el caso general sigue el mismo procedimiento y se cumplirá que $\|g^{(N)} - \widetilde{g}^{(N)}\|_{\square} = \|\widehat{f}^{(N)} - \widetilde{f}^{(N)}\|_{\square,k-1} = o_k(1)$.

Ahora por otro lado, si queremos averiguar si la sucesión de funciones de carga $\{\lambda^{(N)}\}_{N\in\mathbb{N}}$, cumple con la condición de formas k-lineales, deberíamos calcular a las esperanzas en (6), de la Definición 1.5. Pero de acuerdo a como fué definida cada $\lambda^{(N)}$, en términos de $\nu^{(N)}$, se tiene que dichas esperanzas son idénticas a las de (8) en la Definición 1.6, de donde sabemos por hipótesis que $\{\nu^{(N)}\}_{N\in\mathbb{N}}$, cumple con la condición de formas k-lineales, esto se traduce en que $\{\lambda^{(N)}\}_{N\in\mathbb{N}}$, cumplirá con la condición de formas k-lineales (para funciones de carga).

Ahora empleando el Lema de Conteo (Teorema 1.2), se tiene que

$$\mathbb{E}_{\mathbf{x}\in\mathbb{Z}_N^k}[g_{-1}^{(N)}(\mathbf{x}_{-1})\cdots g_{-k}^{(N)}(\mathbf{x}_{-k})] = \mathbb{E}_{\mathbf{x}\in\mathbb{Z}_N^k}[\widetilde{g}_{-1}^{(N)}(\mathbf{x}_{-1})\cdots \widetilde{g}_{-k}^{(N)}(\mathbf{x}_{-k})] + o(1), \tag{10}$$

[6] Véase la nota 4 al pie, de la página anterior.

que de acuerdo a la definición de $g^{(N)}$, el lado izquierdo es igual a la siguiente esperanza $\mathbb{E}_{\mathbf{x}\in\mathbb{Z}_N^k}[\widehat{f}^{(N)}(\psi_1^{(N)}(\mathbf{x}_{-1}))\cdots\widehat{f}^{(N)}(\psi_k^{(N)}(\mathbf{x}_{-k}))]$, ahora si $u = \psi_1^{(N)}(\mathbf{x}_{-1})(\in \mathbb{Z}_N)$, entonces tendríamos lo siguiente $\psi_j^{(N)}(\mathbf{x}_{-j}) - \psi_1^{(N)}(\mathbf{x}_{-1}) = \psi_j^{(N)}(\mathbf{x}_{-j}) - u = (j-1)(x_1+\cdots+x_k) = (j-1)d$, con $d = x_1+\cdots+x_k$, por lo tanto

$$\mathbb{E}_{\mathbf{x}\in\mathbb{Z}_N^k}[g_{-1}^{(N)}(\mathbf{x}_{-1})\cdots g_{-k}^{(N)}(\mathbf{x}_{-k})] = \mathbb{E}_{u,d\in\mathbb{Z}_N}[\widehat{f}^{(N)}(u)\,\widehat{f}^{(N)}(u+d)\cdots\widehat{f}^{(N)}(u+(k-1)d)],$$

y de igual manera

$$\mathbb{E}_{\mathbf{x}\in\mathbb{Z}_N^k}[\widetilde{g}_{-1}^{(N)}(\mathbf{x}_{-1})\cdots\widetilde{g}_{-k}^{(N)}(\mathbf{x}_{-k})] = \mathbb{E}_{u,d\in\mathbb{Z}_N}[\widetilde{f}^{(N)}(u)\,\widetilde{f}^{(N)}(u+d)\cdots\widetilde{f}^{(N)}(u+(k-1)d)],$$

por lo tanto de (10), se sigue que

$$\begin{aligned}&\mathbb{E}_{u,d\in\mathbb{Z}_N}[\widehat{f}^{(N)}(u)\,\widehat{f}^{(N)}(u+d)\cdots\widehat{f}^{(N)}(u+(k-1)d)]\\ &\quad= \mathbb{E}_{u,d\in\mathbb{Z}_N}[\widetilde{f}^{(N)}(u)\,\widetilde{f}^{(N)}(u+d)\cdots\widetilde{f}^{(N)}(u+(k-1)d)] + o(1) \geqslant c(k,\delta) - o_{k,\delta}(1),\end{aligned}$$

esta última desigualdad por el Teorema de Szemerédi (Teorema 0.4), ahora como $\widehat{f}^{(N)} = M_N f^{(N)} \leqslant f^{(N)}$, para todo $N \in \mathbb{N}$, entonces

$$\mathbb{E}_{u,d\in\mathbb{Z}_N}[f^{(N)}(u)\,f^{(N)}(u+d)\cdots f^{(N)}(u+(k-1)d)] \geqslant c(k,\delta) - o_{k,\delta}(1)$$

y esto concluye la demostración. ∎

Con el Teorema Relativo de Szemerédi se concluye esta sección. En la siguiente sección se concentrará en encontrar una función f adecuada que tenga su soporte en el conjunto de los números primos y un conjunto que los contenga y/o una función ν que mayore a f y cumpla con la condición de formas lineales y poder emplear al Teorema Relativo de Szemerédi.

2 | El Teorema de Green-Tao

En esta segunda sección se presentan a los "ingredientes" para poder usar el Teorema Relativo de Szemerédi en el caso específico del conjunto de los números primos, que es precisamente el Teorema de Green-Tao. Los ingredientes que faltan es, una función con soporte en los números primos adecuada, y adecuada en el sentido de que tenga densidad positiva o en términos de la función, que sus esperanzas estén acotadas uniformemente por debajo, por un número positivo, y una función que la acote por arriba y que cumpla con la condición de formas k-lineales (para k fijo).

El *W-trick* y una función mayorante

Una manera de hallar una función que represente a los primos con las características señaladas, sería primero considerar a la función *von Mangoldt* $\Lambda : \mathbb{N} \to \mathbb{R}$, que está definida como

$$\Lambda(n) = \begin{cases} \ln(p) & \text{si } n = p^k, \text{ con algún } p \text{ primo, y } k \in \mathbb{N}, \\ 0 & \text{en otro caso,} \end{cases} \qquad \forall n \in \mathbb{N}.$$

Si pretendemos tomar a esta función como candidata para modelar a los números primos surgen dos complicaciones, una de ellas se puede resolver fácilmente y es que esta función además de considerar a los números primos, también se consideran a sus potencias, así que tendríamos que redefinir Λ para que nada más tenga por soporte a los números primos, lo cual se hará a continuación. El segundo inconveniente surge debido a que los números primos no son tan aleatorios como se desean, pues por ejemplo si $q > 1$ es un entero, los números primos sólo se distribuyen en aquellas clases residuales $a \in \mathbb{Z}_q$, donde $(a, q) = 1$, y recordemos que necesitamos que los números primos estén contenidos en un subconjunto pseudoaletorio con densidad positiva, esto se traduce que necesitamos que la función Λ esté acotada por una función que cumpla la condición de formas k-lineales y en base a esta condición Λ (o los números primos) debería poder distribuirse de manera uniforme sobre cada clase residual. Pero como no está presente esta pseudoaleatoriedad en los números primos, no va a ser posible encontrar una función con estas características que acote a Λ. Entonces como tal sobre todo el conjunto de números primos no va a ser posible definir una función que les asigne peso para usar el Teorema Relativo de Szemerédi.

Lo que proponen Green y Tao es lo que llaman el *W-trick*, y lo que consiguen con este arreglo (en palabras de Terence Tao [9]) es que se aísla a la parte aleatoria de la parte estructurada (no aleatoria) de los números primos. Ahora una de las características que hace a los números primos no tan aleatorios, son estas preferencias a ciertas clases residuales, que sin bien no es posible evitar estas tendencias en general, pero si en cierta medida que involucra al parámetro k que hayamos fijado, de manera que en la condición de formas k-lineales sean imperceptibles (con funciones de error $o_k(1)$) estas tendencias. O sea que al conjunto de los números primos se hará suficientemente aleatorio o pseudoaletorio para que sea posible comprobar la condición de formas k-lineales. En concreto modifican a la función von Mangoldt de la siguiente manera. Si $w : \mathbb{N} \to \mathbb{N}$, es cualquier función que tiende a infinito, suficientemente lento[7] con respecto de N, entonces si W_N es el producto de todos los primos menores o iguales

[7] Es suficiente con pedir que $w(N) \leqslant \frac{1}{2}\ln(\ln(N))$, y $w(N) \to \infty$, con $N \to \infty$, y en base a esto se tendrá que $W_N = O(\ln^{1/2}(N))$.

a $w(N)$, es decir

$$W_N = \prod_{\substack{p \leq w(N) \\ p \text{ primo}}} p, \qquad \forall N \in \mathbb{N}, \tag{11}$$

se define a $\widetilde{\Lambda}^{(N)} : \mathbb{Z}_N \to \mathbb{R}$, como

$$\widetilde{\Lambda}^{(N)}(n) = \begin{cases} \dfrac{\phi(W_N)}{W_N} \ln(nW_N + 1) & \text{cuando } nW_N + 1 \text{ es primo,} \\ 0 & \text{en otro caso,} \end{cases} \qquad \forall n \in \mathbb{Z}_N, \tag{12}$$

para cada $N \in \mathbb{N}$, donde ϕ es la función de Euler.
El factor $\frac{\phi(W_N)}{W_N}$ se utiliza para normalizar, pues el Teorema del Número Primo para Progresiones Aritméticas[8] implica que $\frac{1}{N}\sum_{n=0}^{N-1} \widetilde{\Lambda}^{(N)}(n) = \mathbb{E}_{n \in \mathbb{Z}_N}[\widetilde{\Lambda}^{(N)}(n)] = 1 + o(1)$.
Se afirma que el subconjunto $P_{W_N} = \{n \in \mathbb{N} | \, nW_N + 1 \text{ es primo}\}$ (y a la vez la función $\widetilde{\Lambda}^{(N)}$) no posee ninguna preferencia sobre alguna clase residual de $\mathbb{Z}_q$ siempre que $q < w(N)$, esto puede verse de la siguiente manera, dado que el conjunto de los números primos se distribuye de manera uniforme[8] en particular sobre las clases $\{1, W_N+1, 2W_N+1, \ldots, (q-1)W_N+1\} \subseteq \mathbb{Z}_{qW_N}$, pues $(nW_N + 1, qW_N) = 1$ siempre que $n \in \mathbb{Z}_q$, entonces el conjunto P_{W_N} se distribuirá de manera equitativa sobre todas las clases de $\mathbb{Z}_q$ con $q < w(N)$, aunque todavía habra tendencias para algunas clases de módulo mayor, pero a medida que la función w tienda a infinito con respecto a N, el efecto que produzcan estas tendencias irá disminuyendo. Así que el conjunto P_{W_N} será nuestro candidato para "insertarlo" en un subconjunto pseudoaletorio de los naturales con densidad positiva. En términos de funciones, a la colección $\{\widetilde{\Lambda}^{(N)}\}_{N \in \mathbb{N}}$ les será posible construir una función (colección) $\nu^{(N)} : \mathbb{Z}_N \to [0, \infty)$ que mayore a $\widetilde{\Lambda}^{(N)}$ y que cumpla la condición de formas k-lineales (Definición 1.6).
Como última observación acerca de las funciones $\widetilde{\Lambda}^{(N)}$, notemos que se han descartado a las potencias de los primos y se concentrarán en determinar si es que existe una k-progresión aritmética en el conjunto P_{W_N} (para algún N suficientemente grande), lo que a su vez inducirá una k-progresión aritmética sobre el conjunto de números primos.
Restaría construir a la colección $\{\nu^{(N)}\}_{N \in \mathbb{N}}$ que mayore a $\{\widetilde{\Lambda}^{(N)}\}_{N \in \mathbb{N}}$ y cumpla la condición de formas k-lineales. Como primer intento se puede considerar a $\nu^{(N)} = \widetilde{\Lambda}^{(N)}$, pero la dificultad de probar que $\{\widetilde{\Lambda}^{(N)}\}_{N \in \mathbb{N}}$ cumple con la condición de formas k-lineales es comparable a probar la conjetura de Hardy-Littlewood de las k-tuplas en números primos, por lo tanto se descarta esta posibilidad.
Para poder encontrar una función adecuada que mayore a $\widetilde{\Lambda}^{(N)}$, regresemos a la función von Mangoldt, que a partir de la siguiente identidad conocida

$$\Lambda(n) = \sum_{d|n} \mu(d) \ln\left(\frac{n}{d}\right) = -\sum_{d|n} \mu(d) \ln(d), \qquad \forall n \in \mathbb{N}, \tag{13}$$

[8] **Teorema del Número Primo para Progresiones Aritméticas:** Si $(a, b) = 1$, entonces

$$\sum_{\substack{p \equiv a \mod b \\ p \leq x}} \ln(p) = \frac{x}{\phi(b)}(1 + o(1)),$$

donde aquí $o(1)$, es una cantidad que tiende a cero, cuando x tiende a infinito.

donde $\mu : \mathbb{N} \to \{-1, 0, 1\}$ es la función de *Möbius* definida como $\mu(1) = 1$, y para $n \neq 1$ se define $\mu(n) = (-1)^{\omega(n)}$ si n es libre de cuadrados, donde $\omega(n)$ es la cantidad de números primos (distintos) que dividen a n, y $\mu(n) = 0$, si n no es libre de cuadrados.
Green y Tao en su artículo [4], modifican a la función von Mangoldt a partir de (13), y truncan la suma tal y como ya lo habían hecho Goldston, Pintz y Yildrim en [13], donde investigan la cercania entre primos consecutivos; y la modifican de la siguiente manera, si $R > 0$ se define $\Lambda_R : \mathbb{N} \to \mathbb{R}$ como

$$\Lambda_R(n) = \sum_{\substack{d|n \\ d \leqslant R}} \mu(d) \ln\Big(\frac{R}{d}\Big) = \sum_{d|n} \mu(d) \ln_+\Big(\frac{R}{d}\Big), \qquad \forall n \in \mathbb{N},$$

donde $\ln_+(x) = \max\{\ln(x), 0\}$. No es difícil comprobar que si $1 \leqslant n \leqslant R$ entonces $\Lambda_R(n) = \Lambda(n)$. Mientras que si n no tiene divisores menores o iguales a R entonces $\Lambda_R(n) = \ln(R)$. En este sentido se puede pensar que Λ_R mayoriza a los primos $p > R$, y de hecho también a los *casi primos*, es decir enteros que no tienen divisores menores que R.
En la demostración original de Green y Tao, para su mayorante usan a Λ_R dentro de su expresión (Definición 9.3 [4]). Sin embargo ellos mismos en 2008 [5], ofrecen una prueba simplificada, donde emplean una variante de la función Λ_R, y cambian la reestricción $d \leqslant R$ (lo que es equivalente a $\ln(d)/\ln(R) \leqslant 1$) por una función suave (de clase C^∞) y que tenga soporte[9] compacto, y la proponen de la siguiente manera.

Definición 2.1 Sea $R > 0$. Si $\chi : \mathbb{R} \to \mathbb{R}$, es de clase C^∞, con soporte compacto, se define $\Lambda_{\chi,R} : \mathbb{N} \to \mathbb{R}$, por

$$\Lambda_{\chi,R}(n) = \ln(R) \sum_{d|n} \mu(d) \chi\Big(\frac{\ln d}{\ln R}\Big), \qquad \forall n \in \mathbb{N}. \tag{14}$$

Notemos que recuperamos a la función von Mangoldt si ponemos $\chi(x) = -x$, que aunque χ es suave, no tiene soporte compacto. También podemos recuperar a Λ_R definiendo esta vez a $\chi(x) = \max\{1 - |x|, 0\}$ y en este caso χ tiene soporte compacto, y es continua, pero no es suave.
En general observemos que si $\operatorname{sop}(\chi) \subseteq [-1, 1]$ y $\chi(0) = 1$, entonces χ hará que sólo se consideren divisores de n que sean a lo más R. Observe que en caso de que n no tenga divisores menores o iguales a R, sólo subsistirá en (14) el término $d = 1$, entonces $\Lambda_{\chi,R}(n) = \ln(R)$. Esta propiedad será muy importante para mayorar a $\widetilde{\Lambda}^{(N)}$ sobre los primos.
La simplificación que se hablaba de cambiar a $\ln_+(R/d)$ por $\chi(\ln d/\ln R)$, es en la siguiente dirección, debido a que se más adelante se necesitará estimar sumas que involucran términos al cuadrado de $\Lambda_{\chi,R}$, llega un momento en que conviene expresar a χ como la transformada de Fourier de alguna función, y a la integral que resulte, se le reestringe a un intervalo acotado y aprovechar la suavidad de χ para hacer despreciable a la parte no acotada de la integral en la transformada de Fourier. Así que las condiciones que se le piden a χ son por cuestiones técnicas, además se puede considerar a $\chi(\ln d/\ln R)$ como una aproximación suave a $\ln_+(R/d)$, y de igual manera $\Lambda_{\chi,R}$ se verá como una aproximación a Λ_R.
La siguiente estimación es fundamental, para de ahí partir y poder definir la colección de funciones mayorantes, a partir de la función $\Lambda_{\chi,R}$.

[9] El soporte de una función $f : \mathbb{R} \to \mathbb{R}$, es el conjunto $\operatorname{sop}(f) = \overline{\{x \in \mathbb{R} | f(x) \neq 0\}}$.

Proposición 2.1 *Si fijamos $\chi : \mathbb{R} \to \mathbb{R}$ una función de clase C^∞, con soporte contenido en $[-1, 1]$. Sean $m, t \in \mathbb{N}$ y $\psi_1, \ldots, \psi_m : \mathbb{Z}^t \to \mathbb{Z}$, funciones lineales, donde cualesquiera dos no son multiplos una de la otra.*
Sean $R_N = R(N) = o(N^{1/10^m})$, y $w : \mathbb{N} \to \mathbb{N}$ una función que crece suficientemente lento[10] que depende de N, si W_N está dado por (11); a partir de esto se definen $\theta_i^{(N)} = W_N \psi_i + 1$, para cada $i \in [m]$ y $N \in \mathbb{N}$.
Si $B^{(N)} = \prod_{i=1}^t I_i^{(N)} \subseteq \mathbb{Z}^t$, donde $I_i^{(N)} \subseteq \mathbb{Z}$ es un conjunto con al menos R_N^{10m} enteros consecutivos, entonces se cumple que

$$\mathbb{E}_{\mathbf{x}\in B^{(N)}}\left[\Lambda_{\chi,R_N}\left(\theta_1^{(N)}(\mathbf{x})\right)^2 \cdots \Lambda_{\chi,R_N}\left(\theta_m^{(N)}(\mathbf{x})\right)^2\right] = (1+o(1))\left(\frac{W_N}{\phi(W_N)} c_\chi \ln(R_N)\right)^m \tag{15}$$

donde $c_\chi = \int_0^\infty \chi'(x)^2 dx$, y $o(1)$ denota una cantidad que tiende a cero, cuando N tiende a infinito, y la elección de esta función puede depender de m, t y las funciones $\chi, \psi_1, \ldots, \psi_m$, R y w.
□

En esta estimación es donde se vuelve relevante la definición de W_N, es decir el W-trick, y la cuestión técnica de pedir que w crezca suficientemente lento[10]. La demostración a esta proposición no se expondrá aquí, al ser en cierta medida técnica y extensa, pero al lector interesado se le indica que una prueba puede encontrarse en [1] y en [14], así como en las notas aclaratorias de Tao [10]. Y también se invita a comparar la demostración de la proposición 9.5 de [4], que es la versión (original) de Ben Green y Terence Tao a esta estimación, donde no se emplea en la modificación a Λ, el uso de χ.
Ahora teniendo en cuenta a (15), es que se definirán a las funciones $\{\nu^{(N)}\}_{N\in\mathbb{N}}$, que son el último ingrediente que nos faltaba.

Definición 2.2 (Función mayorante) Fijemos $k \in \mathbb{N}$ con $k \geqslant 3$, y $\chi : \mathbb{R} \to \mathbb{R}$ una función de clase C^∞, con soporte contenido en $[-1, 1]$ y que además $\chi(0) = 1$. Para cada $N \in \mathbb{N}$, sea $R_N = N^{1/k2^{k+3}}$. Si $w : \mathbb{N} \to \mathbb{N}$, en una función que tiende a infinito, suficientemente lento[10] respecto a N, y sea W_N como en (11). A partir de esto se define $\nu^{(N)} : \mathbb{Z}_N \to [0, \infty)$ por

$$\nu^{(N)}(n) = \begin{cases} \dfrac{\phi(W_N)}{W_N} \dfrac{\Lambda_{\chi,R_N}(nW_N+1)^2}{c_\chi \ln(R_N)} & \text{si } n \in [N/2, N), \\ 1 & \text{si } n \in [0, N/2), \end{cases} \qquad \forall n \in \mathbb{Z}_N, \tag{16}$$

y cada $N \in \mathbb{N}$, donde $c_\chi = \displaystyle\int_0^\infty \chi'(x)^2 dx$.

Se eleva al cudrado la función Λ_{χ,R_N}, debido a que no necesariamente es positiva, y se requiere que las funciones $\nu^{(N)}$ sean no negativas.

Prueba del Teorema de Green-Tao

En esta segunda parte de la sección se comprueba que la sucesión $\{\nu^{(N)}\}_{N\in\mathbb{N}}$ cumpla la condición de formas k-lineales, y que mayora a las funciones $\{\widetilde{\Lambda}^{(N)}\}_{N\in\mathbb{N}}$, y así tener válidadas las hipótesis del Teorema de Green-Tao, y poder exponer su demostración.

[10] Véase la nota 7 al pie, de la Página 13.

Proposición 2.2 *La colección de funciones* $\{\nu^{(N)}\}_{N\in\mathbb{N}}$, *definidas en* (16), *cumple con la condición de formas k-lineales, con* $k \geqslant 3$.

Demostración: Fijemos $k \geqslant 3$. Debemos probar que $\{\nu^{(N)}\}_{N\in\mathbb{N}}$, cumple con (8), de la Definición 1.6. Esto es que si $\mathbf{x} = (x_1^{(0)}, x_1^{(1)}, \ldots, x_k^{(0)}, x_k^{(1)}) \in \mathbb{Z}_N^{2k}$, entonces para cada $j \in [\![k]\!]$ y $\omega \in \{0,1\}^k$, se definen $f_{j,\omega} : \mathbb{Z}_N^{2k} \to \mathbb{Z}_N$ por $f_{j,\omega}(\mathbf{x}) = \sum_{i=1}^{k}(j-i)x_i^{(\omega_i)}$, donde $\omega = (\omega_1, \ldots, \omega_k)$ con $\omega_i \in \{0,1\}$. Para cada $N \in \mathbb{N}$ sea $\mathfrak{L}^{(N)} = \{f_{j,\omega} : \mathbb{Z}_N^{2k} \to \mathbb{Z}_N | j \in [\![k]\!],\ \omega \in \{0,1\}^k\}$ el conjunto de estas funciones lineales. Nótese que cualesquiera dos funciones en $\mathfrak{L}^{(N)}$ no son multiplos una de la otra. Entonces la condición de formas k-lineales es verificar que

$$\mathbb{E}_{\mathbf{x}\in\mathbb{Z}_N^{2k}}\left[\nu^{(N)}(\psi_1(\mathbf{x}))\cdots\nu^{(N)}(\psi_m(\mathbf{x}))\right] = 1 + o_k(1), \tag{17}$$

con $1 \leqslant m \leqslant k2^{k-1}$ y $\{\psi_1, \ldots, \psi_m\} \subseteq \mathfrak{L}^{(N)}$.
Ahora como $\nu^{(N)}$ está definida en dos partes no podemos usar la Proposición 2.1 directamente, hay que dividir a $\mathbb{Z}_N$ en intervalos. Sea $Q_N = Q(N)$ una función $(Q : \mathbb{N} \to \mathbb{N})$ que crece lentamente (más adelante se especifica) que depende de N. Ahora se divide a $\mathbb{Z}_N$ en Q_N intervalos más o menos iguales, para formar una partición de $\mathbb{Z}_N^{2k}$ en Q_N^{2k} cajas, en específico para cada colección de $2k$ elementos (no necesariamente distintos) $u_1, \ldots, u_{2k} \in \mathbb{Z}_{Q_N}$, se denotará

$$B^{(N)}_{u_1,\ldots,u_{2k}} = \prod_{j=1}^{2k}\left(\left[u_j\tfrac{N}{Q_N}, (u_j+1)\tfrac{N}{Q_N}\right) \cap \mathbb{Z}_N\right) \subseteq \mathbb{Z}_N^{2k}.$$

Entonces la esperanza en (17), la podemos reescribir de la siguiente forma

$$(1+o_k(1))\mathbb{E}_{u_1,\ldots,u_{2k}\in\mathbb{Z}_{Q_N}}\left[\mathbb{E}_{\mathbf{x}\in B^{(N)}_{u_1,\ldots,u_{2k}}}\left[\nu^{(N)}(\psi_1(\mathbf{x}))\cdots\nu^{(N)}(\psi_m(\mathbf{x}))\right]\right],$$

la función de error $o_k(1)$ surge debido a que los intervalos no son todos del mismo tamaño, y por consecuencia las cajas no todas tienen la misma cantidad de elementos.
Se dirá que una caja $B^{(N)}_{u_1,\ldots,u_{2k}}$ es *buena*, si para cada $j \in [\![m]\!]$, la imagen $\psi_j(B^{(N)}_{u_1,\ldots,u_{2k}})$ está totalmente contenida en el intervalo $[N/2, N) \subseteq \mathbb{Z}_N$ ó completamente fuera de este. De otra forma se dirá que la caja $B^{(N)}_{u_1,\ldots,u_{2k}}$ es *mala*.
Ahora se puede pedir que la función Q tienda a infinito, suficientemente lento de manera que $N/Q_N \geqslant R_N^{10m}$ (donde recordemos que se toma $R_N = N^{1/k2^{k+3}}$, en la Definición 2.2). Así por la Proposición 2.1 y la definición de $\nu^{(N)}$, se tiene que para cajas buenas

$$\mathbb{E}_{\mathbf{x}\in B^{(N)}_{u_1,\ldots,u_{2k}}}\left[\nu^{(N)}(\psi_1(\mathbf{x}))\cdots\nu^{(N)}(\psi_m(\mathbf{x}))\right] = 1 + o_k(1),$$

pues para cajas buenas $\nu^{(N)}(\psi_i(\mathbf{x}))$ es básicamente $\frac{\phi(W_N)}{W_N}\frac{\Lambda_{\chi,R_N}(W_N\cdot n+1)^2}{c_\chi \ln(R_N)}$ ó 1.
Por otro lado, para cajas malas se emplea la cota $\nu^{(N)}(n) \leqslant 1 + \frac{\phi(W_N)}{W_N}\frac{\Lambda_{\chi,R_N}(W_N\cdot n+1)^2}{c_\chi \ln(R_N)}$, así que si se desarrollan los factores, y usando (15) en cada uno de los 2^m términos que surgen, se obtendrá que

$$\mathbb{E}_{\mathbf{x}\in B^{(N)}_{u_1,\ldots,u_{2k}}}\left[\nu^{(N)}(\psi_1(\mathbf{x}))\cdots\nu^{(N)}(\psi_m(\mathbf{x}))\right] = O_k(1),$$

pues de hecho está acotado por $2^m + o_k(1)$.

Concluirá la demostración si se prueba que la proporción de las $2k$-tuplas $(u_1,\dots,u_{2k}) \in \mathbb{Z}_{Q_N}^{2k}$, que producen cajas malas $B^{(N)}_{u_1,\dots,u_{2k}}$, son a lo más $o_k(1)$. En efecto, supongamos que $(u_1,\dots,u_{2k}) \in \mathbb{Z}_{Q_N}^{2k}$ genera una caja mala $B^{(N)}_{u_1,\dots,u_{2k}}$, esto quiere decir que existe $i \in [\![m]\!]$ tal que $\psi_i(B^{(N)}_{u_1,\dots,u_{2k}})$ intersecta tanto a $[N/2, N)$ como a $[0, N/2)$ en $\mathbb{Z}_N$. Esto implica que existe algún[11] $\mathbf{x_0} \in \prod_{j=1}^{2k}[u_j\frac{N}{Q_N}, (u_j+1)\frac{N}{Q_N}) \subseteq (\mathbb{R}/N\mathbb{Z})^{2k}$ con $\psi_i(\mathbf{x_0}) = 0$ ó $\psi_i(\mathbf{x_0}) = N/2$ mód N. Tomando $\mathbf{y_0} = \mathbf{x_0}(Q_N/N)$, se observa que $\mathbf{y_0} \in \prod_{j=1}^{2k}[u_j, u_j+1) \subseteq (\mathbb{R}/Q_N\mathbb{Z})^{2k}$. En este caso $\psi_i(\mathbf{y_0}) = 0$ ó $\psi_i(\mathbf{y_0}) = Q_N/2 \mod Q_N$, y como la longitud de cada intervalo $[u_j, u_j+1) \subseteq \mathbb{R}/Q_N\mathbb{Z}$ es de 1, se tiene que $\psi_i(u_1,\dots,u_{2k}) = \psi_i(\mathbf{y_0}) + O_k(1)$, es decir

$$\psi_i(u_1,\dots,u_{2k}) = O_k(1), \quad \text{ó} \quad \psi_i(u_1,\dots,u_{2k}) = \tfrac{Q_N}{2} + O_k(1) \mod Q_N. \tag{18}$$

Como $\psi_i(u_1,\dots,u_{2k})$ es en general de la forma $\sum_{j=1}^{2k} L_{ij}u_j$, y debido a que ningún ψ_i es la función lineal idénticamente cero, se tendrá que la cantidad de $2k$-tuplas $(u_1,\dots,u_{2k})$ que satisfacen (18), son a lo más $O_k(Q_N^{2k-1})$, pues a lo más podemos elegir $2k-1$ entradas de $(u_1,\dots,u_{2k})$ para poder satisfacer (18), teniendo en cuenta que sólo hay una cantidad finita de funciones lineales, se tiene entonces que la proporción de $2k$-tuplas $(u_1,\dots,u_{2k})$ que generan cajas malas es igual a $O_k(1/Q_N) = o_k(1)$, como se quería probar. ∎

Proposición 2.3 *Para cada $k \in \mathbb{N}$ con $k \geqslant 3$ existen $\delta_k > 0$ y $N_0 \in \mathbb{N}$, tal que $\delta_k\widetilde{\Lambda}^{(N)}(n) \leqslant \nu^{(N)}(n)$ para $\frac{N}{2} \leqslant n < N$, siempre que $N \geqslant N_0$.*

Demostración: Sea $k \geqslant 3$. Empleando la notación de la Definición 2.2, consideremos $\delta_k = (k2^{k+4}c_\chi)^{-1}$, ahora debido al lento crecimiento de la función $w : \mathbb{N} \to \mathbb{N}$, con respecto de N, se observa que $(NW_N+1)/N^2 = o(1)$. A partir de esto existe $N_0 \in \mathbb{N}$ tal que se cumple $\ln(N) \geqslant \ln(NW_N+1)/2$, para $N \geqslant N_0$.
Ahora probemos que $\delta_k\widetilde{\Lambda}^{(N)}(n) \leqslant \nu^{(N)}(n)$ para $\frac{N}{2} \leqslant n < N$, siempre que $N \geqslant N_0$. Podemos suponer que nW_N+1 es primo, pues en caso contrario la desigualdad es trivial, ahora notemos que

$$\ln(R_N) = \ln(N^{1/k2^{k+3}}) = \frac{\ln(N)}{k2^{k+3}} \geqslant \frac{\ln(NW_N+1)}{k2^{k+4}} = c_\chi\delta_k\ln(NW_N+1),$$

para $N \geqslant N_0$.
Como nW_N+1 es primo y $R_N < nW_N+1$, se tiene que $\Lambda_{\chi,R_N}(nW_N+1) = \ln(R_N)$, por lo tanto si $\frac{N}{2} \leqslant n < N$, y de la desigualdad obtenida se cumple que

$$\begin{aligned}\delta_k\widetilde{\Lambda}^{(N)}(n) &= \delta_k\frac{\phi(W_N)}{W_N}\ln(nW_N+1) \leqslant \delta_k\frac{\phi(W_N)}{W_N}\ln(NW_N+1)\\ &\leqslant \frac{\phi(W_N)}{W_N}\frac{\ln(R_N)}{c_\chi} = \nu^{(N)}(n),\end{aligned}$$

como se quería probar. ∎

Y por último el Teorema de Green y Tao que se escribió en la introducción.

[11] Aquí $\mathbb{R}/N\mathbb{Z}$, debe entenderse como el espacio cociente producido por la relación de equivalencia siguiente: si $u, v \in \mathbb{R}$ entonces $u \sim v$ si y sólo si $u - v \in \mathbb{Z}$ y $N|u-v$.

Teorema 0.2 (de Green-Tao)
El conjunto de los números primos contiene progresiones aritméticas de longitud arbitraria.

Demostración: Fijando $k \geqslant 3$, probemos que sobre el conjunto de números primos hay una k-progresión aritmética. Para esto se define para cada $N \in \mathbb{N}$, a la función $f^{(N)} : \mathbb{Z}_N \to [0, \infty)$ como

$$f^{(N)}(n) = \begin{cases} \delta_k \widetilde{\Lambda}^{(N)}(n) & \text{si } n \in [\frac{N}{2}, N), \\ 0 & \text{si } n \in [0, \frac{N}{2}), \end{cases} \qquad \forall n \in \mathbb{Z}_N.$$

Por el Teorema del Número Primo para Progresiones Aritméticas[12] se tendrá que se cumple $\sum_{n\in\mathbb{Z}_N} f^{(N)}(n) = \sum_{N/2\leqslant n<N} f^{(N)}(n) = (\frac{1}{2} + o_k(1))\delta_k N$, luego $\mathbb{E}_{n\in\mathbb{Z}_N}[f^{(N)}(n)] \geqslant \delta_k/3$, con N suficientemente grande. Como $0 \leqslant f^{(N)} \leqslant \nu^{(N)}$ con N suficientemente grande y $\{\nu^{(N)}\}_{N\in\mathbb{N}}$ cumple con la condición de formas k-lineales, se tiene por el Teorema Relativo de Szemerédi que

$$\mathbb{E}_{x,d\in\mathbb{Z}_N}\big[f^{(N)}(x) f^{(N)}(x+d)\cdots f^{(N)}(x+(k-1)d)\big] \geqslant c(k, \delta_k/3) - o_{k,\delta_k}(1).$$

El lado izquierdo puede reescribirse así

$$\frac{\mathbb{E}_{x\in\mathbb{Z}_N}[(f^{(N)}(x))^k]}{N} + \mathbb{E}_{x\in\mathbb{Z}_N, d\in\mathbb{Z}_N\setminus\{0\}}\big[f^{(N)}(x) f^{(N)}(x+d)\cdots f^{(N)}(x+(k-1)d)\big],$$

y se puede observar que $\mathbb{E}_{x\in\mathbb{Z}_N}[(f^{(N)}(x))^k]/N = O_{k,\delta_k}\big(\ln^k(N)/N\big) = o_{k,\delta_k}(1)$. Por lo tanto

$$\mathbb{E}_{x\in\mathbb{Z}_N, d\in\mathbb{Z}_N\setminus\{0\}}\big[f^{(N)}(x) f^{(N)}(x+d)\cdots f^{(N)}(x+(k-1)d)\big] \geqslant c(k, \delta_k/3) - o_{k,\delta_k}(1),$$

debe pasar entonces que para algún número natural $N_k \in \mathbb{N}$ suficientemente grande, se tenga que $f^{(N_k)}(x_k) f^{(N_k)}(x_k + d_k)\cdots f^{(N_k)}(x_k + (k-1)d_k) > 0$, para algún $x_k \in [N_k/2, N_k)$, y un $d_k \neq 0$. Como $f^{(N_k)}$ se anula fuera de $[N_k/2, N_k)$, esto implica que $x_k + (j-1)d_k \in [N_k/2, N_k)$, para cada $j \in [k]$. Por lo tanto $\{x_k + (j-1)d_k\}_{j=1}^k$ es de hecho una k-progresión aritmética en $\mathbb{Z}$, y por la definición de $f^{(N_k)}$ a partir de $\widetilde{\Lambda}^{(N_k)}$, se tendrá que $\{(x_k + (j-1)d_k)W_{N_k} + 1\}_{j=1}^k$ es una k-progresión aritmética de números primos. ∎

[12] Véase la nota 8 al pie, de la Página 14.

3 | Conclusiones

Una vez más los números primos nos revelan lo misteriosos que son. En este caso el Teorema de Green-Tao está como evidencia, pues a pesar de que se tiene que usar de algún modo la cierta aleatoriedad que existe en los números primos, la conclusión del Teorema de Green-Tao nos apunta a que de hecho hay cierta estructura para que dentro de ellos exista una k-progresión aritmética para cada valor de k. Lo cual hasta nos resultaría contradictorio. Pues por un extremo tenemos que justificar esta aleatoriedad y por el otro obtenemos una estructura a nuestro conjunto, y esto se nos hace algo dificil de concebir en un conjunto que en principio no está "construido" de manera natural, o al menos eso es lo que se piensa.

Hablando del Teorema de Green-Tao, está de más, decir que es un resultado monumental, por diversos motivos. Uno de los que primero nos viene a la mente es, que por si sólo el enunciado del teorema, es realmente fácil de entender, pero es de esos teoremas que son (muy) difíciles de probar. Que involucra desarrollar resultados prácticamente igual de difíciles de probar, uno de ellos por ejemplo es una extensión de un teorema muy importante en combinatoria aritmética y nos estamos refiriendo al Teorema de Szemerédi, y es la piedra angular para la demostración de nuestro teorema en cuestión.

Otra de las cualidades que se requiere destacar al Teorema de Green-Tao, ó más bien a su prueba, es la idea de esta pseudoaleatoriedad. La manera en que Green y Tao pronostican que se vea reflejada sobre lo que llaman la condición de formas lineales. Pero también se debe mencionar que si bien son este tipo de ideas que lo hacen un teorema complicado, más no así la teoría que se requiere para ello. Pues no se necesita una sofisticada teoría matemática en particular, ni siquiera de la misma teoría (analítica) de números, pues *a priori* sólo se requiere al Teorema del Número Primo.

Como en la mayoría de los resultados más famosos en la matemática, existen más de una demostración a ellos, no se puede decir de que alguna demostración es mejor a que otra, pero si se puede discutir de las ventajas o desventajas con respecto a otra demostración. Así que intentaremos explicar los aspectos que nos dejó la demostración de Conlon, Fox y Zhao del Teorema de Green-Tao que consideramos como ventajas sobre la prueba original (de Green y Tao en [4]). En principio se pudo observar que la versión de Conlon, Fox y Zhao requiere una cantidad menor de resultados auxiliares que la prueba original. Esto no podría significar tanto si es que dichos resultados hubiesen sido más complejos o haber requerido una teoría más robusta. Pero este no fue el caso, pues la diferencia radica en que Conlon, Fox y Zhao expresan en particular el Lema de Conteo (Teorema 1.2) y prueban el Teorema Relativo de Szemerédi usando la Teoría de Hipergráficas, aunque quizas se está exagerando pues, en el fondo sólo emplean alguna notación de esta, más no algún resultado directo proveniente de esta teoría en particular, y esto representa una ventaja más, pues no se necesita el conocimiento de esta área para poder seguir la demostración.

Por último, aunque en esta tesina no se pudieron ofrecer las pruebas para el Teorema del Modelo Denso y el Lema de Conteo podemos advertir que las demostraciones llevadas a cabo por Conlon, Fox y Zhao en [1] y [2], requieren el desarrollo de menos material para su prueba, con respecto a su contraparte en el trabajo de Ben Green y Terence Tao en [4]. Mientras que los resultados de la segunda sección en esta tesina, prácticamente transcurren de la misma forma, con respecto a la prueba original y sólo representó algunos cambios el uso de la función "moldeadora" χ, sobre la función von Mangoldt modificada Λ_R.

Agradecimientos

Deseo agradecer a la Universidad Nacional Autonoma de México por la oportunidad que me dió de poder realizar mis estudios de posgrado en la maestría de ciencias matemáticas, y por supuesto extiendo mis agradecimientos a la unidad de posgrado de matemáticas por haberme aceptado en su programa, así como al Consejo Nacional de Ciencia y Tecnología por otorgarme una beca para la realización de dichos estudios.

Por supuesto también quisiera agradecer a todos mis profesores a lo largo de mi carrera como estudiante, y en particular a los profesores con los que tuve la oportunidad de convivir y asistir a sus cursos, también quiero agradecer a mi tutora la doctora Magali Louise Marie Folch Gabayet por todos sus consejos, además de haberme orientado a lo largo de mi estancia en esta institución, y al doctor Ernesto Rosales Gonzales quien me brindo diversos consejos para poder realizar esta tesina, y un agradecimiento especial al doctor Timothy Gendron Thornton, por haber aceptado a dirigir, contribuir con sus consejos y correcciones a este trabajo, así como de antemano a la persona encargada de evaluar y emitir su juicio por esta tesina, muchas gracias.

Por último y no menos importante quiero agradecer a mi familia que en todo momento han estado conmigo, me han apoyado, y no valdría la pena nada en el mundo sin ellos.

Referencias

[1] D. Conlon, J. Fox, Y. Zhao, ***The Green-Tao Theorem: An Exposition***, EMS Surv. Math. Sci. (2014), 1-26.

[2] _____, ***A Relative Szemerédi Theorem***, Geom. Funct. Anal. (2015), 1-20.

[3] E. Szemerédi, ***On Sets of Integers Containing no k Elements in Arithmetic Progression***, Acta Arith.(1975), 299–345.

[4] B. Green, T. Tao, ***The Primes Contain Arbitrarily Long Arithmetic Progressions***, Ann. of Math. (2006), 1-56.

[5] _____, ***Linear Equations in Primes***, Ann. of Math. (2008), 1-84.

[6] H. Furstenberg, ***Ergodic Behavior of Diagonal Measures and a Theorem of Szemerédi on Arithmetic Progressions***, Journal d'Analyse Mathématique, (1977), 204–256.

[7] P. Varnavides, ***On Certain Sets of Positive Density***, J. London Math. Soc. (1959), 53–97.

[8] T. Tao, ***A Quantitaive Ergodic Theory Proof of Szemerédi Theorem***, Electronic J. Combinatorics, (2004), 1-52.

[9] _____, ***The Dichotomy Between Structure and Randomness, Arithmetic Progressions, and The Primes***, (2006), 581-608.

[10] _____, ***A Remark on Goldston-Yildirim Correlation Estimates***, 1-8.

[11] Y. Zhao, ***An Arithmetic Transference Proof of a Relative Szemerédi Theorem***, Math. Proc. Cambridge Philos. Soc. (2014), 255–261.

[12] O. Reingold, L. Trevisan, M. Tulsiani, S. Vadhan, ***New Proofs of the Green-Tao-Ziegler Dense Model Theorem: An Exposition***, (2008), 1-7.

[13] D.A. Goldston, J. Pintz, C.Y. Yildirim, ***Small Gaps Between Primes***, (2005), 1-34.

[14] T. Bloom, ***The Green-Tao Theorem on Arithmetic Progressions Within The Primes***, (2010), 1-60.

[15] G.H Hardy, E.M. Wright, ***An Introduction to The Theory of Numbers***, 6^a edición Oxford University Press. (2008), 1-621.