

# PROYECTO FIN DE CARRERA

Presentado a

LA UNIVERSIDAD DE LOS ANDES  
FACULTAD DE CIENCIAS  
DEPARTAMENTO DE FÍSICA

Para obtener el título de

**FISICO**

por

*Camilo Andrés Rivera Lozano*

---

***IMPACTO DE LOS PARÁMETROS COSMOLÓGICOS EN LA  
ESTRUCTURA A GRAN ESCALA DEL UNIVERSO***

---

Sustentado el – de Diciembre de 2014 frente al jurado:

## Composición del jurado

- *Asesor:* Jaime E. Forero-Romero PhD, Profesor Asociado, Universidad de Los Andes

# Tabla de contenido

<b>1. Introducción</b>	<b>1</b>
1.1. Motivación . . . . .	2
1.2. Objetivos . . . . .	2
<b>2. Marco Teórico</b>	<b>3</b>
2.1. La evolución del universo . . . . .	3
2.1.1. El Principio cosmológico y el problema del Horizonte . . . . .	3
2.1.2. La Era Inflacionaria . . . . .	4
2.1.3. La inestabilidad gravitacional y la composición del universo . . . . .	4
2.1.4. Variables Relevantes y Parametrizaciones . . . . .	6
2.1.5. Formación de estructura y halos de materia oscura . . . . .	7
2.2. Simulación en Paralelo y Análisis . . . . .	9
2.2.1. Gadget-2 . . . . .	9
<b>3. Metodología del Trabajo</b>	<b>12</b>
3.1. Cronograma de Trabajo . . . . .	13
<b>4. Trabajo Realizado</b>	<b>14</b>
4.1. Características de Simulación . . . . .	14
4.1.1. Parámetros de instalación y compilación . . . . .	14
4.1.2. Parámetros de la simulación . . . . .	15
4.1.3. Simulaciones realizadas . . . . .	16
4.2. Códigos de Análisis . . . . .	18
4.2.1. Estructura del <i>Snapshot</i> . . . . .	18
4.2.2. Extracción de Halos . . . . .	18
4.2.3. Campo de densidades . . . . .	19
4.3. Análisis y Resultados . . . . .	20
<b>5. Conclusiones</b>	<b>22</b>
5.1. Trabajo Futuro . . . . .	22
<b>Bibliografía</b>	<b>22</b>

# Índice de figuras

2.1. Falso vacío y verdadero vacío . . . . .	4
2.2. Diagrama de tiempo del universo . . . . .	7
2.3. División del espacio con octree . . . . .	10
2.4. Ejemplos de Curvas de Peano-Hilbert . . . . .	11
3.1. Metodología de trabajo . . . . .	12
4.1. Dos simulaciones de $128^3$ partículas con iguales condiciones a excepción de $\sigma_8$ . . . . .	16
4.2. Comparación de tamaños entre simulaciones de prueba y simulaciones definitivas . . . . .	17
4.3. Distribución de cambios en la velocidad para $\Omega - 5\%$ . . . . .	20
4.4. Distribución de cambios en la velocidad para $\Omega + 5\%$ . . . . .	20
4.5. Distribución de cambios en la masa para $\Omega - 5\%$ . . . . .	21
4.6. Distribución de cambios en la masa para $\Omega + 5\%$ . . . . .	21

# Índice de tablas

3.1. Cronograma de Trabajo . . . . .	13
4.1. Parámetros de simulaciones definitivas . . . . .	18

# 1. Introducción

El universo, su composición y evolución han sido objeto de estudio para el hombre desde tiempos incluso anteriores a poseer la capacidad para ver más allá del cielo. La innegable inmensidad del universo nos deja sin más remedio que la observación para su caracterización y, siendo lo más rápido del universo, la luz es por excelencia el método a través del cual podemos ahondar en los secretos que esconde.

El creciente desarrollo de la tecnología, la mejora en la sensibilidad de los equipos y la capacidad de computación para el análisis de los datos nos han brindado un mejor panorama de cómo es el universo, tanto local como a gran escala, pero el hecho que finalmente nos mostró la verdad sobre la historia y antigüedad del universo fue el descubrimiento y análisis del fondo de radiación de microondas. No sólo fue el sustento para la teoría del *Big-Bang* sino que también nos mostró como fue el universo en momentos cercanos a su origen, justo después de la recombinación cuando el universo se volvió transparente y la luz pudo finalmente viajar libremente.

Estos análisis nos permitieron conocer a fondo la composición del universo y caracterizarlo a partir de los que hoy se conocen como los parámetros cosmológicos, pero ¿qué tan diferente sería el universo si variaran dichos parámetros? Lamentablemente contamos con un único universo y el método de la experimentación es inviable pues las condiciones en las que se generó el universo exceden por mucho nuestras capacidades, de manera que la única alternativa para responder esta pregunta yace en las simulaciones computacionales a gran escala.

La creación de máquinas de muy alto nivel de procesamiento y la paralelización de tareas a lo largo de un gran número de procesadores ( $> 50$ ) ha permitido la creación de simulaciones cada vez más grandes en un lapso de tiempo cada vez menor. Las simulaciones de  $N$  cuerpos permiten observar la evolución del universo y extraer información de las posibles diferencias en cada uno de los escenarios propuestos. En la actualidad se encuentran en desarrollo diferentes experimentos, entre otros el DESI[3] y el HETDEX[5], que pretenden medir y cuantificar los efectos de la energía oscura en el universo, de manera que los resultados de este proyecto podrán contrastarse con los datos arrojados por estos experimentos mostrando la distribución de energía-materia que más se acopla a nuestro universo.

El siguiente documento presenta la planeación, desarrollo y resultados de un trabajo que consistió en la generación de condiciones iniciales para universos con diferentes parámetros cosmológicos, la simulación en paralelo de la evolución de dichos universos y el análisis de los estados finales de cada simulación. Para esto se hizo uso de la paralelización del trabajo en cajas con condiciones periódicas de entre 150Mpc y 500Mpc, un número de partículas entre  $128^3$  y  $512^3$  y un tiempo de simulación cercano a los 13Gyr. El análisis realizado a cada estado final fue hecho a través del lenguaje C, mientras que la visualización de los datos fue realizada a través de Python.

En el Capítulo 2 se explicarán las bases teóricas y los conocimientos pertinentes que fueron requeridos para el desarrollo del proyecto, desde el funcionamiento de una simulación en paralelo, la historia del universo y la descripción de los parámetros cosmológicos. Posteriormente en el Capítulo 3 se explicará brevemente la metodología planteada y usada a través del desarrollo del proyecto, el plan de

trabajo y los pasos que se siguieron. Luego, en el Capítulo 4, se mostrarán los resultados del trabajo realizado, así como los problemas encontrados, los códigos y programas creados y la intercomunicación entre estos.

## 1.1. Motivación

Las mediciones más recientes de las anisotropías del universo primitivo, grabadas en el CMB, fueron realizadas por la sonda Planck en donde se establecieron los valores más actualizados de los parámetros cosmológicos. De todas maneras, aún quedan muchas incógnitas por resolver, y son grandes cambios los que se pueden alojar dentro del margen de error de las mediciones. Este proyecto trabaja desde el marco de las simulaciones, realizando variaciones del 5% en la densidad de materia en el universo alrededor de las mediciones de la sonda Planck, para identificar los rasgos que caracterizan dichas variaciones. En especial se centra en las diferencias observables de la formación y dinámica de los halos de materia oscura, como lo son sus velocidades de centro de masa y su abundancia en el universo. Estos datos enriquecerán los datos de las mediciones y correcciones realizadas sobre los parámetros cosmológicos, permitiendo un análisis más profundo en donde se puede correlacionar las mediciones con las observaciones de efectos indirectos.

## 1.2. Objetivos

### General

Cuantificar el cambio de la estructura a gran escala del universo ante escenarios con diferentes parámetros cosmológicos.

### Específicos

- Obtener una serie de universos simulados ante diferentes valores de parámetros cosmológicos.
- Extraer información acerca de los diferentes universos como la abundancia de halos de materia oscura y las distribuciones de velocidad entre pares de halos.
- Realizar un análisis comparativo entre los diferentes universos simulados

## 2. Marco Teórico

### 2.1. La evolución del universo

El universo tal y como lo conocemos hoy está poblado de cuerpos colosales, sorprendentes e incluso hermosos. Las estrellas, las nebulas y las galaxias son formaciones a las que estamos acostumbrados y que están presentes en donde quiera que nos adentremos en el cielo. Pero fue un proceso muy largo y turbulento para llegar a lo que hoy observamos y consideramos habitual, como nuestro sistema solar, un proceso que aún está lleno de incógnitas y de teorías que si bien se ajustan a las observaciones y predicciones, no están completamente comprobadas. La siguiente sección trata de resumir el conjunto de teorías y hechos que hacen parte de la visión más ampliamente aceptada como la historia del universo.

#### 2.1.1. El Principio cosmológico y el problema del Horizonte

Las teorías que describen los momentos que precedieron al *Big-Bang* y el análisis del fondo de radiación de microondas (CMB por sus siglas en inglés) han brindado un panorama de lo que parece ser el inicio del universo y su progreso hasta lo que conocemos hoy en día; un universo que cumple con el principio cosmológico[7]. Las fórmulas de la relatividad general de Einstein muestran cómo la estructura del Espacio-Tiempo se ve curvada por la presencia de la masa, pero al intentar aplicar estas fórmulas para mostrar la dinámica del universo, hay que considerar un modelo para éste. Dada la complejidad matemática que pueden llegar a tener estas fórmulas en el caso general, fue necesario considerar el caso más simple posible, dictado por el principio cosmológico. Bajo este principio, el universo es isotrópico y homogéneo, lo cual quiere decir que el universo tiene una distribución homogénea y esta distribución se cumple en cualquier dirección en la cual se observe. Se podría pensar que los planetas, estrellas, galaxias y otros cuerpos nombrados anteriormente, los cuales además de todo están separados por vastos vacíos casi absolutos, serían el perfecto contraejemplo para la suposición de Einstein, pero este enunciado hace referencias a escalas mucho mayores a las que podría abarcar cualquiera de estos cuerpos. Para dar una idea de la escala de distancias tratadas, la Vía Láctea tiene un tamaño aproximado de  $30kpc$ , pero si partimos el universo en cubos, por ejemplo de  $90Mpc$ , veremos que estos cubos comparten características casi idénticas, como por ejemplo, la densidad promedio, emisión de luz e incluso temperatura [4].

Este hecho antes de ser tranquilizante (vivimos en lo que sería el caso ideal de un universo), es bastante alarmante puesto que no puede ser coincidencia que lugares tan apartados del universo posean la misma temperatura, y no son dos lugares sino todos los rincones observados del universo los que cumplen con dicho principio. El problema radica en que la forma de comunicación más rápida existente entre dos cuerpos es la luz, y por ende la comunicación entre dos lugares del universo estará limitada por la distancia que los separa. De esta manera, si consideramos que el universo tiene una edad aproximada de  $13,8Gyr$ , no hay forma para que dos puntos opuestos (separados por  $\approx 28Gly$ ) hayan tenido algún intercambio tal que les hubiese permitido estar en condiciones iguales hoy en día.

### 2.1.2. La Era Inflacionaria

La teoría que hasta la fecha mejor resuelve esta incongruencia conocida como el *Problema del Horizonte* es la teoría de la inflación. Ésta supone que en sus primeros instantes (al rededor de tiempos de  $\sim 10^{-33} \text{seg}$ ), el universo pasó por un proceso de crecimiento acelerado, llegando incluso a velocidades superlumínicas. Existen varias versiones de esta época y no se ha logrado explicar a fondo la física detrás de este comportamiento. Entre las formas de explicarlo o describirlo, una explica que esta expansión se dio gracias a un estado que se conoce como *falso vacío*, el cual no es mínimo estado de la densidad energética para un campo escalar, pero sí es un mínimo local, sólo que es inestable. En dicho estado se pudo dar una inflación de una sección tan pequeña que pudiese ser homogénea en sí, causando que el espacio resultante de la inflación replicara estas características. Dado el carácter atractivo (signo negativo) de la fuerza gravitacional es posible una expansión en donde se conserve la energía (Se “crea” energía al expandir que es compensada por la *energía negativa* de la gravedad)[4].

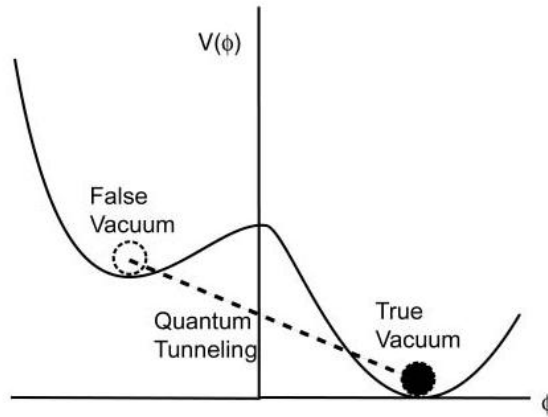


Figura 2.1: Falso vacío y verdadero vacío <sup>1</sup>

Otra forma de ver la inflación es pensando en que cualquier irregularidad presente en el universo cuando aún cabía en la palma de la mano se vería alisada por una expansión de dichas magnitudes. La difusión de los primeros fotones logró erradicar cualquier irregularidad en pequeña y gran escala de toda la materia ligada a la radiación electromagnética[7]. En cualquiera de los casos, el resultado final de la inflación es un universo con las características que observamos en todas direcciones.

Se cree que las galaxias y cuerpos que actualmente pueblan el universo se generaron a raíz de las fluctuaciones cuánticas presentes en él cuando su tamaño fue lo suficientemente pequeño para que éstas tuvieran un efecto relevante en su estructura. En ese caso estas fluctuaciones de pequeña escala crecerían hasta tener un gran tamaño y se convertirían con el paso del tiempo en lo que hoy observamos como cuerpos estelares. Pero *¿cómo es que estas fluctuaciones cuánticas siguieron presentes si la época inflacionaria borró cualquier irregularidad dando a un universo homogéneo?* La respuesta reside en cerca de un cuarto del contenido energético del universo: la materia oscura.

### 2.1.3. La inestabilidad gravitacional y la composición del universo

Observaciones y mediciones de los espectros de galaxias y estrellas lejanas no sólo nos brindan información de su composición, sino que también de su dinámica. Las alteraciones en el espectro pueden

<sup>1</sup>Imagen tomada de [http://ned.ipac.caltech.edu/level5/Watson/Watson5\\_3.html](http://ned.ipac.caltech.edu/level5/Watson/Watson5_3.html)



deberse a varios factores, en especial el corrimiento hacia el rojo, para el cual existen tres principales motivos. El primero es debido al Efecto Doppler, que al igual que con el sonido, la longitud de onda aumenta (se corre hacia el rojo) cuando el emisor y el observador se alejan a grandes velocidades; el segundo es dictado por la relatividad general, cuando una onda electromagnética está escapando de un campo gravitacional; y el tercero, de mayor interés para efectos de la discusión a seguir, es el corrimiento al rojo (o *redshift*) intrínseco debido a la expansión del universo.

Observaciones de diferentes espectros atómicos en el universo han mostrado que incluso corrigiendo los corrimientos dados por movimientos relativos y por campos gravitacionales, siempre hay un corrimiento al rojo remanente. Un fotón emitido en un dicho momento tendrá una longitud de onda definida, pero si el universo se expande (el espacio donde se desplaza se expande), al cabo de un tiempo este fotón tendrá una longitud de onda menor que la que tenía inicialmente. De esta manera se ha observado que el universo se está expandiendo, y además es una herramienta de gran utilidad para definir distancias de objetos lejanos. Al observar objetos distantes, observamos su pasado, pues lo que nos llega es la luz emitida tiempo atrás, cuando el universo era un poco más pequeño, de manera que a mayor redshift en el espectro, más antiguo es el objeto que observamos.

Si consideramos que la densidad es constante en el universo, esto permite realizar un análisis local obviando el efecto del resto del universo dada la simetría esférica. La energía total por unidad de masa en una esfera de radio  $r$  es una expresión dada por la suma de energía cinética, por el movimiento de las partículas, y la energía gravitacional potencial existente entre estas, así,  $E = \frac{1}{2}v^2 - \frac{GM}{r}$ , en donde  $G$  es la constante gravitacional y  $M$  estará dada considerando una densidad  $\rho$  uniforme en el volumen de la esfera. Reorganizando los términos, se puede llegar a la Ecuación (2.1), en donde  $\Omega = \frac{\rho}{\rho_c}$  y el término  $\rho_c$  se conoce como la densidad crítica.

$$E = \frac{1}{2}v^2(1 - \Omega) \quad (2.1)$$

Si la densidad promedio de materia en el universo  $\rho$  es menor que la densidad crítica  $\rho_c$  entonces tendremos que  $E < 0$  lo cual significa que la gravedad nunca podrá contrarrestar la expansión actual y toda galaxia está condenada a una muerte solitaria pues esta expansión continuará indefinidamente. Esto se conoce como la solución de un universo abierto. En cambio, si la densidad está por encima de la densidad crítica se tiene que la expansión se detendrá en un punto y comenzará a revertirse debido a la atracción gravitacional terminando en lo que se ha acuñado como el *Big-Crunch*. Esta segunda solución se conoce como el universo cerrado. Finalmente, la tercera solución corresponde al universo plano en la que  $E = 0$  y la densidad de materia es exactamente igual a la densidad crítica. Esta tercera es la que se ha observado que coincide con la realidad, pero este es un equilibrio críticamente inestable, ya que por más homogéneo que se considere el universo, cualquier fluctuación crecerá con el paso del tiempo. Así si una pequeña sección supera la densidad crítica, comenzará a ejercer una mayor fuerza que el resto del universo, absorbiendo materia de las partes en donde la densidad es menor a la crítica. Rápidamente la parte densa se volverá más y más densa, contrario a lo que sucede con las secciones menos densas; una vez roto el equilibrio no hay vuelta atrás.

Como se dijo en la Sección 2.1.2, la inflación eliminó cualquier fluctuación de materia interactuante con la radiación, de manera que se estableció un equilibrio irrompible, en donde la formación de estrellas y galaxias se vuelve imposible. La materia oscura es la teoría de mayor acogida y validez para romper no sólo esta paradoja, sino que también otros fenómenos observados como las lentes gravitacionales y las velocidades de rotación de sistemas binarios. Para todos estos, se necesita que exista una cantidad de materia mayor que la observada; es decir, una cantidad mayor de materia que la que interactúa con la radiación electromagnética (materia ordinaria). De allí su nombre, la materia oscura es un tipo de materia con las mismas características gravitacionales que la materia ordinaria, pero no interactúa con ningún tipo de luz, de manera que no se puede observar de manera directa. La única forma de

detectarla es por efectos indirectos, como la medición la cantidad de materia necesaria para que los fenómenos nombrados anteriormente tengan un sentido físico.

Volviendo al universo primitivo, las fluctuaciones cuánticas generaron cambios en todo tipo de materia, pero la inflación sólo afectó aquella materia interactuante con la radiación, de manera que estas irregularidades quedaron grabadas en la materia oscura y crecieron junto con todo el universo. Esto creó regiones con mayor atracción gravitacional, causando que la materia ordinaria cayera en dichas zonas, juntándose, creciendo y formando cuerpos celestes. De manera que (bajo este punto de vista) es gracias a la presencia de la materia oscura que se pudo dar un universo tal y como lo conocemos, y es más, simulaciones de muy alto nivel de cuerpos de materia oscura, en comparación con las imágenes de estructura a gran escala del universo, muestran que la materia ordinaria se organiza exactamente en donde se generan mayores concentraciones de materia oscura, por el efecto nombrado anteriormente. Mediciones de la distribución de materia muestran que la materia oscura suma aproximadamente un 25 % de la energía total del universo, mientras que la materia ordinaria forma alrededor de un 5 %.

Estos porcentajes se conocen por mediciones y estimaciones, e incluyéndolos en la Ecuación (2.1), puesto que según las observaciones el universo debería ser plano y  $\rho \approx \rho_c$ . Al sumar toda la densidad de materia oscura y materia ordinaria, aún faltaría un gran porcentaje de densidad energética para cumplir la condición nombrada arriba, de manera que se plantea la energía oscura para cumplir con esa cuota. Es muy poco lo que se sabe acerca de la energía oscura y son muchas las teorías y vertientes entorno a ella, desde la energía de polarización de vacío hasta un fluido que llena todo el espacio tiempo. Entre las cosas que se pueden afirmar de ésta es que es cerca del 70 % del contenido energético del universo y es la responsable de la expansión de este (contrarrestada únicamente por la materia que conforma el otro 30 %).

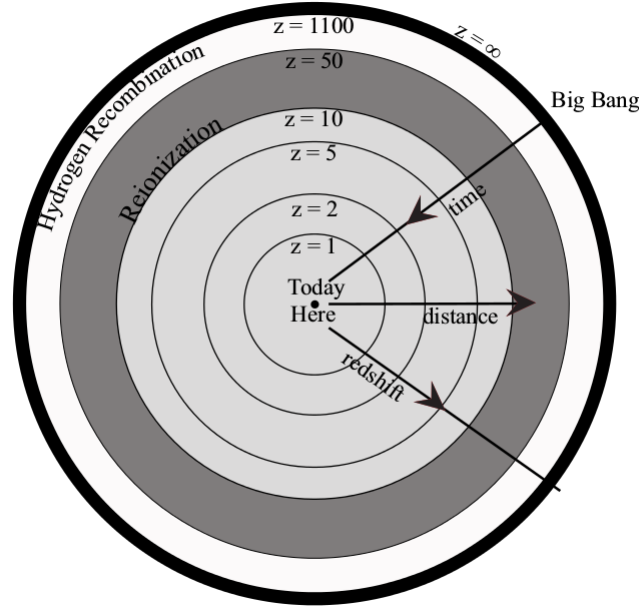
#### 2.1.4. Variables Relevantes y Parametrizaciones

Una vez explicada la historia que conllevó a la evolución del universo, es importante rescatar las parametrizaciones de cada fenómeno nombrado anteriormente, puesto que estas variables son los parámetros de ingreso a las simulaciones que se realizaron. Empezando por la expansión del universo, esta se mide por un factor de escala  $a(t)$  que varía con el tiempo, tiene un valor de 1 en la actualidad y para su definición se dice que si una distancia tiene una longitud  $r(t_1)$  en dicho tiempo, entonces en  $t_2$  estará definido por,

$$r(t_2) = r(t_1) \left[ \frac{a(t_2)}{a(t_1)} \right] \quad (2.2)$$

Este factor es el que permite llevar cuenta de la expansión del universo a lo largo de una simulación, y puede definirse a partir del *redshift*  $z$ ,  $a(t) = \frac{1}{1+z(t)}$  del cual se hablaba en la Sección 2.1.3. En la Figura 2.2 se observa un diagrama que muestra la historia del universo desde el Big-Bang, en la parte más externa cuando  $z$  tiende a infinito, hasta la actualidad en el centro, donde  $z = 0$ . Bajo este parámetro se fijan los tiempos de simulación y los intervalos de generación de *snapshots* para las simulaciones realizadas. Asimismo, debemos incluir la versión matemática de la ley de expansión de Hubble, en donde se establece (a partir de observaciones) que la velocidad relativa (y por ende el corrimiento al rojo) de un objeto distante será proporcional a su distancia, donde  $H_0$  es conocida como la constante de Hubble,

$$v = H_0 r \quad (2.3)$$

Figura 2.2: Diagrama de tiempo del universo<sup>2</sup>

Pero la constante de Hubble puede cambiar con el avance del tiempo, de manera que como las demás constantes se define un valor actual y su avance hacia atrás en el tiempo. Por ejemplo, volviendo a la ecuación 2.1, se asume que en el presente  $\Omega_0 = 1 = (\Omega_\Lambda + \Omega_m + \Omega_r)$ , para lo cual la primera contribución corresponde a la densidad de energía del vacío (Energía Oscura), el segundo término corresponde a la materia (tanto oscura como ordinaria o bariónica ( $\Omega_b$ ) y la contribución energética de la radiación. Estos parámetros de distribución de energía son las semillas que pueden llevar a diferentes universos y son precisamente los parámetros de entrada para las simulaciones realizadas en este proyecto. La expresión general para los parámetros está dada por la Ecuación (2.4).

$$\frac{H(t)}{H_0} = \left[ \frac{\Omega_m}{a^3} + \Omega_\Lambda + \frac{\Omega_r}{a^4} \right]^{\frac{1}{2}} \quad (2.4)$$

### 2.1.5. Formación de estructura y halos de materia oscura

Para describir el origen de la Ecuación (2.4) y la dinámica del universo en general, se vuelve necesario introducir una serie de variables que puedan llevar cuenta de los cambios existentes en él. Las primeras son conocidas como las coordenadas comóviles, las cuales a diferencia de las coordenadas fijas tienen en cuenta el factor de expansión del universo  $a(t)$ , de manera que todo vector de posición  $\vec{r}$  tiene asociado unas coordenadas comóviles  $\vec{x} = \vec{r}/a$ . Por otro lado, al describir cualquier magnitud vectorial es necesario introducir un marco de referencia; por ejemplo, en la tierra el marco de referencia estático es el planeta en sí mismo y el movimiento de los objetos es referido respecto al suelo. Pero cuando comenzamos a trabajar en un panorama más amplio, como lo son las grandes secciones del universo, y donde el vacío domina impidiendo dar una referencia clara, es importante introducir lo que se conoce como las velocidades peculiares. Al caracterizar el movimiento y distancia de estrellas y galaxias lejanas,

<sup>2</sup>Imagen tomada de [7]

la velocidad peculiar es aquella componente de velocidad que no se explica a través de la Ley de Hubble de la Ecuación (2.3), y viene siendo la velocidad del marco de reposo local, dada por la Ecuación (2.5).

$$\vec{u} = \vec{v} - H\vec{r} \quad (2.5)$$

Paralelo a esto, se modela la expansión cosmológica como un fluido ideal, sin presión y en expansión. Además, definen las fluctuaciones fraccionales de densidad como se muestra en la Ecuación (2.6), en donde  $\bar{\rho}$  corresponde a la densidad de masa promedio. Finalmente se define el potencial gravitacional a través de la ecuación de Poisson para la gravitación Newtoniana, mostrada en la Ecuación (2.7).

$$\delta(\vec{x}) = \frac{\rho(\vec{r})}{\bar{\rho}} - 1 \quad (2.6)$$

$$\nabla^2 \phi = 4\pi G \bar{\rho} a^2 \delta \quad (2.7)$$

Estas definiciones se realizan para plantear la ecuación de continuidad y la Ecuación de Euler que permiten describir el fluido así:

$$\frac{\partial \delta}{\partial t} + \frac{1}{a} \nabla \cdot [(1 + \delta)\vec{u}] = 0 \quad (2.8)$$

$$\frac{\partial \vec{u}}{\partial t} + H\vec{u} + \frac{1}{a}(\vec{u} \cdot \nabla)\vec{u} = -\frac{1}{a}\nabla\phi \quad (2.9)$$

Todas las ecuaciones mostradas anteriormente dan como resultado la ecuación diferencial mostrada en (2.10). Hay que tener en cuenta que el comportamiento del fluido estará bien definido para la evolución de partículas de materia oscura sin colisión siempre y cuando no se crucen. Pero los cruces se darán una vez crezcan las perturbaciones y la relación sea no lineal con  $|\delta| > 1$ . La Ecuación (2.10) Tiene dos soluciones, de las cuales sólo una es creciente, y es proporcional a  $D(t)$ , el cual es llamado el factor lineal de crecimiento [6]. En épocas tempranas del universo cuando dominaba la materia ( $1 < z < 10^3$ ) tenemos un factor de crecimiento tal que  $D(t) \propto t^{\frac{2}{3}} \propto a(t)$ , pero a medida que avanzamos, los crecimientos son cada vez más lentos, al punto que después de la época inflacionaria, cuando el universo es dominado por la Constante de Hubble, todas las fluctuaciones en el campo de densidades quedan congeladas. Lo que esto quiere decir es que por más que existan lugares en el universo en el que se generan condiciones cerradas que colapsan en sí mismas, i.e. galaxias, el potencial gravitacional permanece intacto en magnitud, y por ello las fluctuaciones que observamos en el potencial gravitacional que observamos hoy en día son del mismo orden de magnitud que las que se observan en el CMB.

$$\frac{\partial^2 \delta}{\partial t^2} + 2H \frac{\partial \delta}{\partial t} = 4\pi G \bar{\rho} \delta \quad (2.10)$$

Cuando una sobredensidad comienza a atrapar material y entra en la zona de crecimiento no lineal, esta zona frena su expansión, hay un punto de retorno y finalmente colapsa. Estas regiones son conocidas como *halos* de materia oscura, y allí es donde la materia ordinaria se combina para generar nuevos cuerpos. Es por eso que toda galaxia está rodeada o inmersa en un halo de materia oscura, lo explica la velocidad de rotación tal que sigue existiendo un estado ligado (si ni hubiera presencia de la materia oscura, la velocidad de rotación sería superior a la velocidad de escape) para las estrellas y sistemas formados. Así que si se modela las regiones de sobredensidad como una región esférica cuya función

de densidad cumple con un perfil cuadrado (“sombrero de copa”), se pueden resolver analíticamente las ecuaciones de evolución llegando a que una región esférica dejará de expandirse para colapsar a un punto, si su sobredensidad en un tiempo  $z$  está dada por,

$$\delta_{crit}(z) = \frac{1,686}{D(z)} \quad (2.11)$$

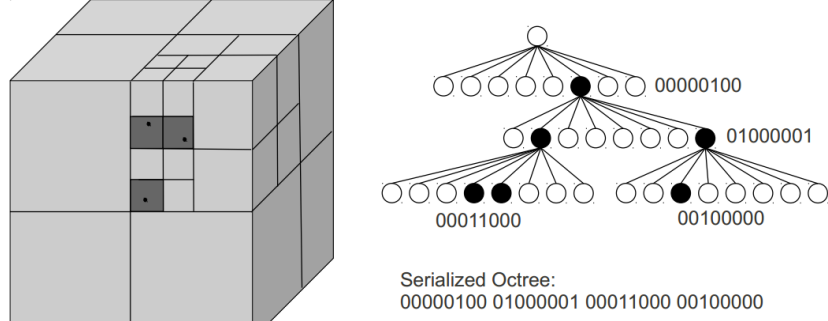
## 2.2. Simulación en Paralelo y Análisis

Como se dijo en la Sección 2.1, la materia oscura tiene las mismas propiedades gravitatorias que la materia ordinaria u observable, pero la supera en cantidad por casi 5 veces en todo el universo. Es por esto que para analizar las estructuras a gran escala el universo, la distribución de materia, los grandes halos y la dinámica de la materia en el universo se puede partir del análisis de materia oscura, puesto que se espera que la materia ordinaria esté ubicada en donde existan mayores concentraciones de materia oscura. Si bien no tenemos forma de observar los cambios en el universo si se cambia la distribución de energías, se pueden realizar simulaciones a gran escala de todo el universo variando los parámetros cosmológicos. En la siguiente sección se explicará brevemente las características de los algoritmos utilizados y su modo de funcionamiento.

### 2.2.1. Gadget-2

Toda simulación de un entorno estará limitada por el recursos computacionales que se tengan, como por ejemplo el tiempo de simulación, la cantidad de procesadores disponibles y la cantidad de memoria que se puede ocupar. De manera que la eficiencia de un código estará dictada por la eficiencia en consumo de dichos recursos; qué tantos recursos consume y qué tan exactos son los resultados obtenidos. Existe una gran variedad de códigos para la simulación de  $N$  partículas con diversas ventajas y desventajas entre cada uno, pero la principal diferencia radica en el método de análisis de cada código para realizar aproximaciones a pequeña y gran escala, permitiendo así una cierta tolerancia sin comprometer mayor cantidad de recursos. Gadget-2 es un código creado por Volker Springel y publicado en el 2005[8] para realizar grandes simulaciones cosmológicas de  $N$  cuerpos. La principal ventaja de este código es la efectividad en el manejo de gran y pequeña escala, pues dentro de un cierto rango, realiza cálculos detallados de la gravedad, pero para mayores distancias realiza un análisis de Fourier que calcula la gravedad efectiva aproximada.

La primera característica del código de Gadget-2 es la creación de un *octree* BH para el análisis de las partículas, lo cual significa que se crea un árbol de análisis en el cual cada nodo puede tener 8 hojas o ramificaciones. Este algoritmo permite que la fuerza gravitacional computacional calculada para cada partícula sea realizada por  $\mathcal{O}(\log N)$  interacciones, contrario a los algoritmos de suma directa los cuales necesitan  $\mathcal{O}(n^2)$  interacciones. Para la generación del *octree* se parte del cubo inicial, el cual se divide en 8 octantes iguales y luego se decide si cada subdivisión debe partirse nuevamente en 8 partes hasta que cada partición realizada (hoja del árbol) tiene 1 o 0 partículas en ella, como se puede observar en la Figura 2.3.

Figura 2.3: División del espacio con octree <sup>3</sup>

De esta manera, al tener construido el árbol, sólo resta recorrerlo para calcular las fuerzas efectivas. Así que se inicia por la raíz del árbol o el nodo principal y de acuerdo a las características del nodo, se decide si se explorará más allá o no, y en dado caso recorrer las 8 ramificaciones para la contribución de la fuerza efectiva. En el caso de que un nodo esté lo suficientemente lejos, toda la sección del árbol que se encuentra de allí en adelante se considera como la contribución de un único cuerpo. Finalmente para la efectividad del código y la compatibilidad con el SPH (*Smoothed particle hydrodynamics*) no se realizan expansiones multipolo del momento sino que se usa el monopolo de momento.

$$\phi(x) = \sum_i m_i \varphi(x - x_i) \xrightarrow{\mathcal{F}} \phi_k^{short} + \phi_k^{long} \quad (2.12)$$

La Ecuación (2.12) muestra el análisis que se debe realizar para el potencial peculiar gravitacional por Gadget-2, pero además de esto, se realiza una descomposición en el espacio de Fourier para discernir entre los efectos lejanos y cercanos. Para la parte real es suficiente realizar un análisis en el espacio real mediante la función de error complementario, como se muestra en la Ecuación (2.13), en donde  $r_i = \min(|\vec{x} - \vec{r}_i| - nL)$ , ya que la función de error complementario rápidamente suprime cualquier contribución distante. Por otro lado, para las contribuciones lejanas del potencial se vuelve mucho más sencillo realizar un análisis mediante métodos de mallas en espacio de Fourier. Este proceso, en conjunto con el método Tree, se le llama frecuentemente como el *TreePM Method*.

$$\phi_k^{short}(x) = -G \sum_i \frac{m_i}{r_i} \text{erfc}\left(\frac{r_i}{2r_s}\right) \quad (2.13)$$

Con este método nunca es necesario recorrer todos los nodos del árbol, pues este recorrido se trunca rápidamente para el espectro corto, y el efecto de las partículas distantes se calcula con ayuda del campo de densidades estimado por el método CIC (*Cloud in Cell* explicado en la Sección 4.2.3) y la transformada rápida de Fourier (FFT). Así se logra optimizar el uso de recursos, puesto que la información de cada partícula puede ser almacenada en el mismo espacio en el que en algún punto estuvo la información de los 8 nodos de conexión y además se logra optimizar las interacciones lejanas por métodos optimizados de Fourier.

Por último, también es importante hablar del sistema de paralelización de Gadget-2 para la división del problema varios procesadores. Dada la cantidad de partículas e interacciones a tener en cuenta en cada simulación se vuelve imperativo distribuir la gran tarea entre múltiples procesadores, de manera que cada cual se encargue de una parte diferente, pero para esto hay múltiples formas de lograrlo. La forma

<sup>3</sup>Imagen tomada de <http://traumabot.blogspot.com/2013/06/octree-based-point-cloud-stream.html>

escogida en Gadget-2 es una arquitectura de código que se pueda correr en cualquier sistema, desde un computador personal, hasta grandes clusters de cientos de procesadores, en donde cada procesador disponible cuenta con su espacio de memoria propio y se encarga de un problema de menor tamaño al original. Esto se logra mediante un código abierto llamado MPI (*Message Passing Interface*) el cual hace uso de las librerías de FFTW y GSL. Una vez escogida la forma de paralelización, aún falta crear un algoritmo que permita dividir el trabajo total en tareas menores, de manera eficiente.

La estrategia usada en este código es la generación de una curva fractal llamada Peano-Hilbert, que permite llenar cualquier dominio con formas simples e iterativas, dependiendo de la cantidad de divisiones que se apliquen. En la Figura 2.4 se pueden ver dos ejemplos de lo que sería la curva para llenar todo el espacio, tanto en dos dimensiones (2.4a) para varias divisiones, como en tres dimensiones (2.4b). Al realizar las divisiones del espacio, de cualquier manera, hay que recordar que cada división hace una parte del todo, y por lo tanto habrá relaciones e intercambio de información y partículas entre cada segmento, y es por esto que la división debe realizarse con mucho cuidado, puesto que entre mejor distribuido el trabajo será más equilibrada la comunicación entre partes. La ventaja de usar las curvas de Peano-Hilbert es que al segmentar la curva en puntos específicos, las regiones resultantes tienen una baja relación de aspecto y relación superficie a volumen, lo cual tiene una alta conveniencia para la reducción de comunicación entre dominios adyacentes.

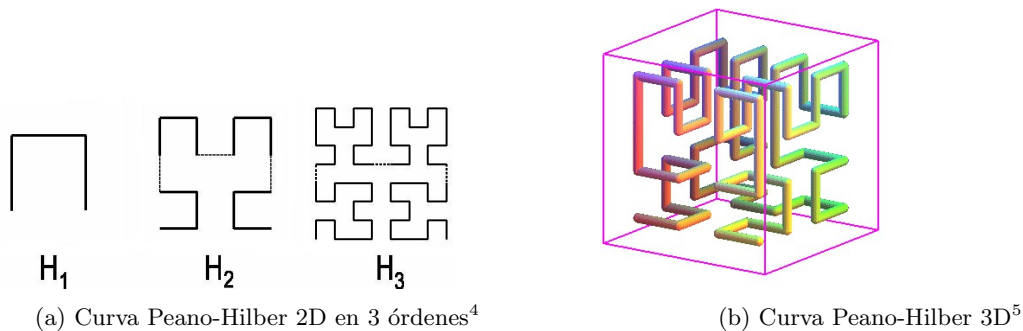


Figura 2.4: Ejemplos de Curvas de Peano-Hilbert

El método de división, como se puede observar en la Figura 2.4, logra que sin importar la división, las partículas de cada dominio serán espacialmente cercanas, y por ende con interacciones similares. Finalmente en cada dominio (en cada procesador) se realiza el algoritmo de TreePM, introduciendo pseudo-partículas que representarán los resultados de dominios adyacentes (u otros procesadores). Se realiza el recorrido completo por el árbol, y en caso de necesitar la información de una pseudo-partícula, se solicita al procesador correspondiente. Una vez los procesadores terminan de realizar el recorrido de su árbol local proceden a mandar dicha información a quienes la solicitaron. La alta correlación que existe en el método TreePM y la división por Peano-Hilbert permite que los errores de cálculo de fuerzas sean independientes del número de procesadores usado y de los lugares en donde se divide la curva.

<sup>4</sup>Imagen tomada de <http://pc-ua.com.ua/page120.html>

<sup>5</sup>Imagen tomada de <http://www.mathcurve.com/fractals/hilbert3d/hilbert3d.shtml>

### 3. Metodología del Trabajo

El desarrollo de este proyecto se basó en una metodología completamente computacional, debido a que el trabajo principal fue la realización de múltiples simulaciones y la creación e implementación de códigos y programas para su análisis. La metodología propuesta y usada a lo largo del semestre se puede resumir en la Figura 3.1 en donde se pueden ver las diferentes etapas del proyecto. En la primera etapa se consultó el material bibliográfico disponible tanto para la generación de simulaciones a través de Gadget-2 [8], así como para la comprensión de los conocimientos básicos para el desarrollo del proyecto, explicados en el Capítulo 2.

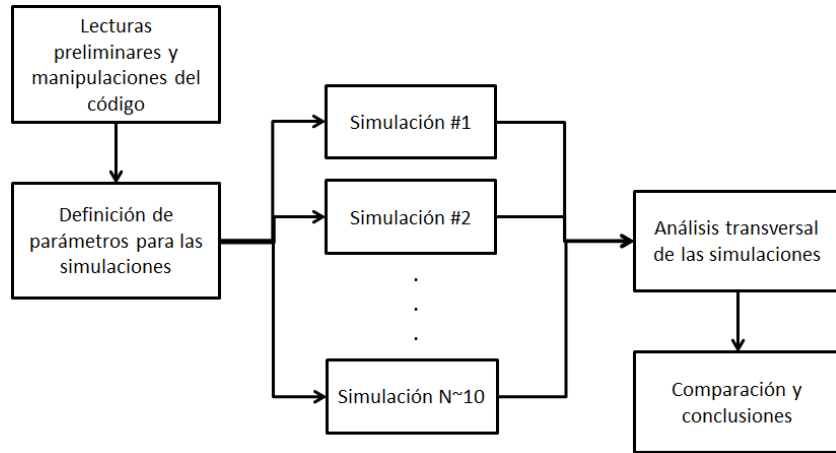


Figura 3.1: Metodología de trabajo

Una vez instaladas todas las librerías y códigos necesarios en el cluster de física, se procedió a realizar una serie de simulaciones preliminares. Éstas tenían como objetivo principal observar la dinámica del código, sus posibles errores, medir el tiempo de computación y sobretodo tener una muestra de los archivos exportados en cada simulación para poder analizarlos, ver su estructura y la información que contenían y en base a ello comenzar a escribir códigos de análisis que servirían para las simulaciones grandes. Estas primeras simulaciones se basaban en cajas cúbicas periódicas de lado de  $150Mpc$  y contenían  $128^3$  y  $256^3$  partículas.

Una vez realizadas todas las pruebas preliminares se prepararon las condiciones iniciales para las simulaciones definitivas, las cuales por restricción de tiempo y recursos computacionales se fijaron a un cubo de  $500Mpc$  de lado con un total de  $512^3$  partículas. Dado el número de partículas, el tiempo de simulación comenzó a aumentar en un factor de 8, así que mientras que las simulaciones de  $128^3$  partículas tomaban alrededor de 3 horas, las siguientes tomaron un día entero. Desafortunadamente al momento de correr las simulaciones definitivas, estas no fallaron por falta de memoria en el cluster



de la universidad, por más que se corrieran en los 56 procesadores no lograba tener registro de todas las partículas.

El análisis realizado a los resultados de las simulaciones tuvo dos vertientes principales: análisis por halos de materia oscura y análisis de fluctuaciones de densidades. De igual manera se generaron dos secuencias de código independientes para cumplir con ambos propósitos. Para la primera se hizo uso de un código existente para la búsqueda de halos por medio del algoritmo FoF y posteriormente se usó la información de los halos, como por ejemplo sus centros de masa y velocidades respectivas, tanto para un análisis intrínseco de cada simulación como para un análisis transversal entre las simulaciones. Para la segunda vertiente se calculó el campo de densidades de cada una de las simulaciones para realizar un análisis del espectro de fluctuaciones en cada caso. Para cada una de las simulaciones finales se varió la proporción de materia y energía oscura en el universo.

### 3.1. Cronograma de Trabajo

A continuación, en la Tabla 3.1 se muestra el plan de trabajo propuesto para la realización del proyecto, junto con la distribución de tiempo establecida para la realización de cada una de las tareas.

Actividad	Duración	Descripción
Lectura de Bases teóricas	2 semanas	Revisión bibliográfica acerca de la composición del universo y los parámetros que definen su evolución.
Lectura e instalación del código	2 semanas	Instalación de librerías y entornos necesarios. Familiarización con los métodos de simulación y los códigos y lenguajes a usar.
Simulaciones preliminares	3 semanas	Realización de simulaciones de bajo nivel, observando archivos exportados y comportamiento.
Simulaciones definitivas	4 semanas	Adquisición de datos definitivos de análisis con simulaciones de alto nivel computacional.
Análisis de simulaciones y creación de códigos	3 semanas	Creación de programas de análisis en base a los formatos de exportación de Gadget-2 y generación de estadísticas y comparaciones.
Redacción de documento	3 semanas	Recoger los datos obtenidos y resultados de los análisis para redactar documento final.

Tabla 3.1: Cronograma de Trabajo <sup>1</sup>

<sup>1</sup>La realización de las tareas no fue secuencial y los tiempos mostrados se sobrelapan durante el semestre

## 4. Trabajo Realizado

El trabajo realizado a lo largo del semestre se puede dividir en dos secciones principales: la realización de simulaciones y la escritura de códigos de análisis y comparación para dichas simulaciones. Para el primer caso implica la lectura de cómo correr las simulaciones en clusters de múltiples procesadores, definir los parámetros de instalación y de simulación en sí y observar los comportamientos, tiempos de duración y modos de exportación de datos de cada simulación. Para la segunda sección se incluyen los códigos realizados y utilizados que permiten desde la lectura de los datos brutos de simulación hasta la extracción de propiedades y generación de gráficas de análisis. A continuación se explicarán las diferentes secciones y detalles del trabajo realizado, así como los resultados de dicho trabajo.

### 4.1. Características de Simulación

Los datos manejados a lo largo del proyecto son 100 % provenientes de simulaciones computacionales, de manera que la correcta planeación, manejo y ejecución de dichas simulaciones es vital para los resultados. A continuación se describen las características de la instalación y ejecución de las simulaciones realizadas.

#### 4.1.1. Parámetros de instalación y compilación

Para correr una simulación en Gadget-2 se requiere principalmente de 3 librerías previamente instaladas y de un set de condiciones iniciales sobre las cuales trabajar. Las tres librerías que se deben instalar son FFTW, GSL y MPI, todas ellas de libre distribución e instalación. FFTW es una librería, o conjunto de librerías, con códigos para la realización de transformadas de Fourier en espacio discreto, como la transformada rápida de Fourier (FFT), cuyo uso se explicó en la Sección 2.2.1. GSL (*GNU Scientific Library*) como su nombre lo dice es una librería con un gran conjunto de funciones de uso científico. Finalmente, MPI (*Message Passing Interface*) es la librería de instrucciones, rutinas y métodos para la paralelización de procesos en múltiples procesadores.

Una vez instaladas las librerías se procede a compilar los ejecutables, para lo cual es necesario configurar el MakeFile. En éste se comentan las características y utilidades que no se requieren usar de Gadget-2, así como se activan las que sí. Para el caso del proyecto, los puntos más importantes del MakeFile son activar las condiciones de frontera periódicas, activar el mallado PM (para el algoritmo TreePM explicado en la Sección 2.2.1), el uso de Peano-Hilbert para la paralelización y desactivar el uso de librería HDF5, la cual no fue instalada. Finalmente, se indica las ubicaciones de las librerías previamente instaladas, lo cual cambia dependiendo de la arquitectura del sistema y se procede a compilar para generar el ejecutable de la simulación. Este proceso es idéntico para Gadget-2 y para N-GenIC, que es el código utilizado para el tercer requerimiento: las condiciones iniciales. N-GenIC es un código de

generación de condiciones iniciales que a partir de un archivo *glass* de partículas homogéneamente distribuidas en un mallado genera un archivo del tipo snapshot de Gadget-2 con una distribución de masas y velocidades homogénea. El archivo *glass* utilizado tenía un total de  $16^3$  partículas de manera que de acuerdo a los requerimientos de la simulación (número total de partículas) se puede apilar una cierta cantidad de archivos en cada eje, o crear uno de mayor tamaño.

Como se explicará más a fondo en la Sección 4.1.3, las simulaciones fueron realizadas no sólo en computadores personales sino que también se hizo uso del clúster de física de la Universidad de los Andes y además se corrieron las simulaciones definitivas en el clúster KIAS en Corea, para lo cual fue necesario la conexión y ejecución de comandos remotos mediante la terminal de Ubuntu. Mediante el comando *ssh*, con un usuario y una clave, se establece una conexión remota con el clúster o máquina en cuestión que permite correr cualquier comando o script. En el caso de la ejecución de las simulaciones es importante tener en cuenta la arquitectura y manejo que tenga cada clúster, puesto que en un clúster de bajo uso como lo era el Fiscluster se podía ejecutar directamente el script de la simulación en los nodos que estuviesen desocupados, siempre y cuando los resultados de simulación quedasen dentro de las limitaciones de espacio de disco duro y computación. En cambio, para un cluster de mayor nivel de organización e infraestructura como lo es KIAS, hay que cumplir con una serie de recursos máximos permitidos y además el script no puede ser ejecutado directamente, sino que debe contener la información de la cantidad de nodos y procesadores requeridos y correr el comando *qsub* con el scrip. Este comando añade el script a un sistema de colas, para el cual el administrador de colas, en base a los requerimientos descritos para el script, asigna los procesadores en los que correrá el proceso. Esta organización permite una mejor distribución de los recursos, evitando la sobrecarga de algún nodo. El estado del proceso puede ser seguido mediante el comando *qstat* el cual muestra los procesos corriendo por el usuario actual y con el comando *ghost* que muestra el estado actual de uso de los recursos del clúster, como la carga de los nodos y la memoria usada.

Dado que en el clúster no se hacía uso de las colas de trabajo y los scripts se corrían directamente desde la consola, en el caso de haber una desconexión del computador y romper el acceso remoto, las simulaciones se detenían. Para corregir este problema se hizo uso del comando *nohup* el cual permitía que el código se siguiese ejecutando sin necesidad de la conexión permanente por consola, así todos los outputs que normalmente aparecerían en la consola se guardaban en un archivo de texto.

#### 4.1.2. Parámetros de la simulación

Una vez se tiene listo el ejecutable, se prepara el script para la ejecución de la simulación, mediante el comando *mpirun* o *mpiexec* los cuales deben recibir como entrada el número de procesadores en los que se va a correr la simulación, un archivo con la lista específica de los nodos a usar, el ejecutable y cualquier entrada que requiera dicho ejecutable. Tanto para el caso de N-GenIC como para Gadget-2 se requiere un archivo de parámetros en el cual se especifican las condiciones de simulación, la ubicación de datos de entrada y datos de salida y los parámetros cosmológicos, entre otros.

Los datos más importantes para tener en cuenta con el archivo de parámetros de N-GenIC son: los parámetros cosmológicos  $\Omega$ ,  $\Omega_\Lambda$ ,  $\Omega_b$ ,  $H$ ,  $\sigma_8$ ; el *redshift* inicial de la simulación,  $z$ ; el tamaño de la caja, en kiloparsecs; el *TileFac*, que corresponde al número de veces que se replica el archivo *glass* en cada dirección (*i.e.* número de partículas); y los parámetros de escritura y lectura, como carpetas y bases para nombres de exportación. El parámetro cosmológico  $\sigma_8$  es el cual determina cómo son las fluctuaciones de densidad en el universo, debido a que da cuenta de qué tan dispersas o desviadas son dichas fluctuaciones en esferas de  $8Mpc$ , de manera que como se puede observar en la Figura 4.1, en una simulación con un  $\sigma_8$  mayor, habrá un mayor contraste (serán más pronunciadas las fluctuaciones) de densidades.

En cuanto al archivo de parámetros de Gadget-2 es un poco más extenso, pero la esencia se conserva, comenzando porque se deben volver a incluir los parámetros cosmológicos y el tamaño de la caja, para lo

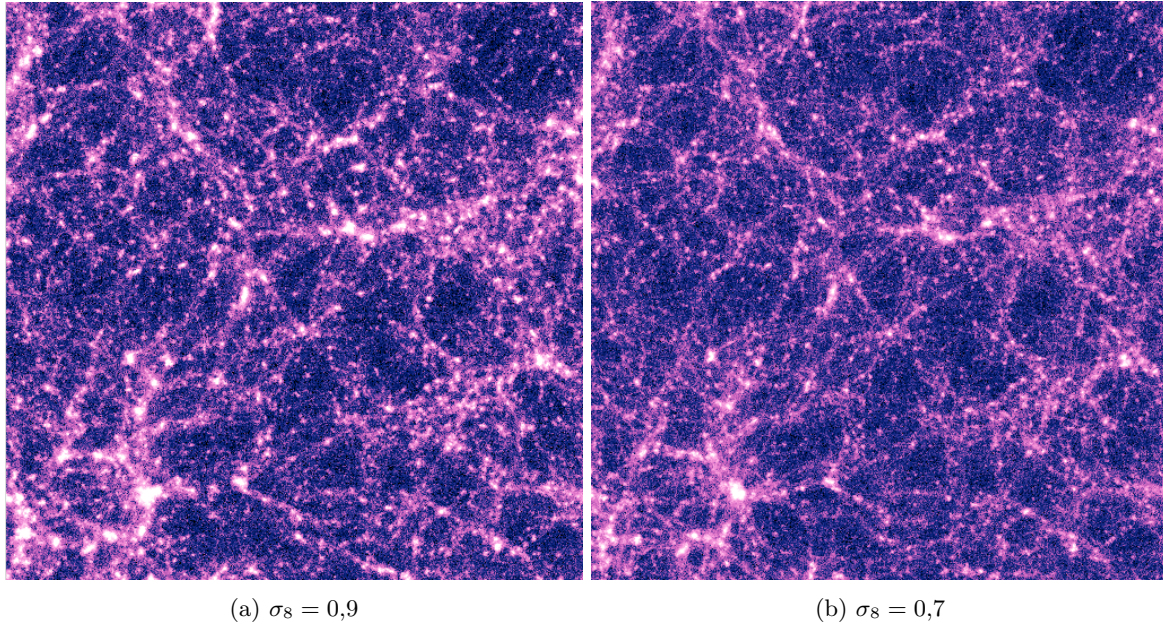


Figura 4.1: Dos simulaciones de  $128^3$  partículas con iguales condiciones a excepción de  $\sigma_8$ . También se aprecian las condiciones periódicas de la simulación.

cual no sobra recordar que dichos valores deben coincidir con los utilizados en el archivo de parámetros de las condiciones iniciales. Nuevamente se deben especificar detalles acerca de la importación de condiciones iniciales (resultados de N-GenIC) y exportación de estados parciales, resultados y archivos de reinicio para cada procesador. Gadget exporta los resultados parciales y totales en archivos de estados donde se tienen todos los datos de la simulación actual en un punto dado; este archivo se explica más a fondo en la Sección 4.2.1. En el archivo de parámetros se deben incluir nuevamente los tiempos de inicio de la simulación, pero esta vez en términos del parámetro de dilatación  $a = \frac{1}{1+z}$  y además el tiempo del primer snapshot y el intervalo de tiempo de cada cuando debe exportarse un snapshot (ya sea por un factor aditivo o multiplicativo). Entre otros parámetros para modificar se encuentran el tiempo máximo de simulación, los errores y tolerancias de los métodos de Tree BH y SPH. Finalmente se tienen las distancias de *softening*, para las que se aconseja usar un décimo del equivalente a la distancia media libre.

### 4.1.3. Simulaciones realizadas

A lo largo del desarrollo del proyecto se corrieron múltiples simulaciones con diferentes características y propósitos, desde las simulaciones de prueba, las cuales permitían observar si el código corría correctamente, hasta las simulaciones definitivas en donde lo más importante eran las variables de entrada. Inicialmente, para la familiarización con el entorno de ejecución y con los modos de ejecución, se corrieron simulaciones de prueba en un computador portátil, lo cual daba para un máximo de dos procesadores. Una vez comprobado que el método de compilación y ejecución de Gadget-2 y N-GenIC era el correcto se procedió a instalar nuevamente las librerías en el clúster del Departamento de Física de la Universidad de los Andes, *fiscluster*. Una vez listas, se corrieron las primeras simulaciones para  $128^3$  partículas, cuyos resultados corresponden a la Figura 4.1. Estas primeras simulaciones fueron realizadas en una caja periódica de tamaño de  $150 Mpc$  y tomaron aproximadamente 3 horas de cómputo en 32 procesadores. Para estas primeras simulaciones, cada snapshot exportado por Gadget-2

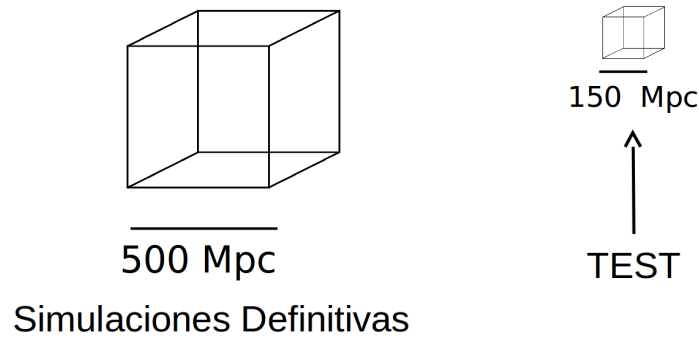


Figura 4.2: Comparación de tamaños entre simulaciones de prueba y simulaciones definitivas

pesaba alrededor de  $57Mb$ . En la Figura 4.2 se puede apreciar una escala del tamaño de la caja en las simulaciones definitivas respecto a las simulaciones de prueba.

El siguiente paso a seguir fue realizar simulaciones de  $256^3$  partículas en una caja periódica de  $500Mpc$  de lado, pero esta vez variando el parámetro de  $\Omega$  de 0,3 a 0,35. El tiempo de simulación tomado para estas simulaciones fue de aproximadamente 24 horas corriendo en 40 procesadores. Debido a la gran cantidad de partículas en la lista, cada snapshot generado para este segundo grupo de simulaciones pesó  $449Mb$ . Para el grupo definitivo de simulaciones realizadas se tomó nuevamente la caja de  $500Mpc$ , pero nuevamente se aumentó el número de partículas en un factor de 8, dando así un total de  $512^3$  partículas; 64 veces más que en las simulaciones de prueba. El tamaño final de cada snapshot generado para esta cantidad de partículas fue de  $3,6Gb$ . Para este último grupo de simulaciones hubo una limitación de memoria física para correr la simulación, incluso corriendo en el máximo de 56 procesadores en fiscluster, no fue posible completar las simulaciones. Fue por esta razón que se decidió migrar la simulación al clúster KIAS en Corea, con una capacidad de 256 procesadores.

El migrar de clúster implicó reinstalar las librerías necesarias para la simulación y también correr nuevamente simulaciones pequeñas para entender el modo de funcionamiento de los trabajos. Este clúster funciona con un sistema de colas de trabajos, de manera que los scripts ser presentados a un administrador de colas quien se encarga de ejecutarlos. Una vez en proceso, se crean dos archivos con radical igual al script, en donde se puede realizar el seguimiento del output de la ejecución (para el primero) y los posibles errores (en el segundo). El cambio de clúster supuso una notoria mejoría en los tiempos de ejecución, puesto que para la creación de las condiciones iniciales se debían realizar en un único procesador, pues Gadget-2 tiende a tener errores al leer un archivo de condiciones iniciales dividido. Así, en fiscluster la creación de estos archivos tomaba alrededor de 20 minutos o más, mientras que en KIAS dicho proceso tomó 5 minutos. Las simulaciones definitivas tomaron aproximadamente 4 días en total, y los parámetros usados para dichas simulaciones se muestran en la Tabla 4.1 y son datos tomados del proyecto Plank.

Parámetro	Valor
$\Omega_0$	$0,3175 \pm 0,05$
$\Omega_\Lambda$	$0,6825 \mp 0,05$
$\Omega_b$	0,048
$H$	0,6711
$\sigma_8$	0,834

Tabla 4.1: Parámetros de simulaciones definitivas

## 4.2. Códigos de Análisis

Una vez realizadas las simulaciones, es necesario crear códigos y programas que permitan la extracción de la información necesaria para los propósitos del proyecto. En la siguiente sección se describen brevemente los códigos creados para dicho propósito. Los algoritmos de análisis siguieron dos ramificaciones principales: el análisis por densidades y el análisis por halos. Para el segundo se tiene un proceso de búsqueda y clasificación de halos para un posterior análisis transversal de las simulaciones. Los lenguajes de programación usados fueron principalmente C para la extracción de información y análisis de la misma y Python (más exactamente la utilidad de IPython notebook) para la generación de gráficas.

### 4.2.1. Estructura del *Snapshot*

Los snapshots de Gadget-2 están escritos en binario, lo cual permite una velocidad de escritura mayor y a su vez un menor consumo de memoria física. Para la lectura de los snapshots se utilizó el código *read\_snapshot.c* el cuál se encuentra en la carpeta de Analysis de la instalación de Gadget-2. Este código permite la lectura el snapshot y la extracción de las posiciones y velocidades de cada partícula en cada eje, de manera que los posteriores códigos tengan acceso a dicha información para su análisis.

El snapshot inicia con un encabezado en el cual indica la cantidad de partículas que hay en cada tipo de los 6 cuerpos manejados por Gadget-2, estos son: gas, halo, disco, *bulge*, estrella y frontera. Los siguientes 6 datos en el encabezado indican la masa asignada a cada tipo de cuerpo, la cual varía debido a que se busca una densidad homogénea y se varía el número de partículas y el tamaño de la caja. Los datos de masa en el snapshot están dados en unidades de  $10^{10} M_\odot$ . A continuación se listan ciertos parámetros del estado de simulación, como lo son el tiempo, el redshift, el tamaño de la caja y los parámetros cosmológicos. Una vez terminado el encabezado, comienza el listado de partículas primero indicando la posición en tres coordenadas de cada una, en unidades de kiloparsec. Una vez se termina el listado de posiciones, inicia el listado de velocidades en tres dimensiones, en *km/s*. Finalmente, se realiza un nuevo listado con el número de identificación ID de cada partícula. El código también permite realizar una conversión de unidades de energía y una reorganización de partículas de acuerdo a si número de identificación.

### 4.2.2. Extracción de Halos

Para la extracción de los halos se hizo uso del código HackFOF el cual utiliza el algoritmo de *Friend of Friends* (FoF) para identificar los halos de materia oscura. Para este algoritmo se define un linking

length, el cual se definió aproximadamente como 0,2 veces la distancia media esperada entre partículas, y se procede a recorrer las partículas, buscando para cada una cuales caen dentro de una esfera de dicho radio. Una vez encontrada la lista que cumplen con dicha condición, se repite el proceso con las partículas en dicha lista, y así hasta que no se agregue ninguna partícula nueva a la lista, y la lista resultante es un halo de materia oscura.

Para la utilización de este código, se debe exportar previamente un archivo de texto con la información de las partículas (posiciones en tres dimensiones separadas por un espacio), precedido por un encabezado similar al del snapshot, en el que se indican el número total de partículas, el número total de partículas tipo halo, el número de partículas tipo gas, el número de partículas tipo estrella, el tiempo y un factor llamado *nactive*, que para efectos del proceso se le asigna cero. Cuando este archivo es exportado se puede proceder a ejecutar el código FoF haciendo uso de las banderas para el uso de la opción de condiciones periódicas.

Una vez terminado de correr el FoF, éste exporta un nuevo archivo que contiene la información de la cantidad total de partículas, seguido por una lista con el índice del halo correspondiente a la partícula en dicha posición; los índices correspondientes a las partículas están en el mismo orden de las partículas ingresadas a HackFOF. El índice cero significa que la partícula no está asignada a ningún halo, mientras que cualquier otro índice indica el número del halo. Este archivo es cargado nuevamente para realizar un análisis del número de halos existentes en la simulación y calcular los centros de masa y velocidades de centro de masa respectivos para cada uno. Para las simulaciones de  $128^3$  partículas se encontraron alrededor de 6000 halos, para la simulación de  $256^3$  aproximadamente 70000 halos y para las simulaciones definitivas de  $512^3$  partículas se encontraron aproximadamente 200000 halos.

### Identificación de correspondencias

Dado que la idea del proyecto es observar cambios en los universos simulados, fue necesario encontrar halos correspondientes entre las simulaciones. Esto es posible ya que en el archivo de parámetros de la generación de condiciones iniciales, uno de los parámetros corresponde a una semilla de generación de números aleatorios y al dejar esta semilla invariante, los universos tendrán estructuras y comportamientos similares. Para la búsqueda de estos halos correspondientes, se realiza un recorrido a través de los halos de la simulación de referencia (aquella con los valores de Planck) y para cada halo se busca dónde están las partículas correspondientes en las demás simulaciones, finalmente se selecciona aquel con mayor número de coincidencias, claro está, siempre y cuando sea un valor razonable.

#### 4.2.3. Campo de densidades

El cálculo del campo de densidades de la simulación y la sobredensidad en cada punto es la segunda vertiente del código, pero que para los efectos finales del proyecto complementa los datos de la primera vertiente. Para el cálculo de las densidades se utiliza el algoritmo conocido como *Cloud in Cell*, el cual, en vez de realizar un conteo bruto de la cantidad de masas presentes en cada división, realiza una ponderación de acuerdo a la posición relativa al centro de las celdas.

El primer paso es dividir la caja en un cierto número de celdas, que serán la resolución del análisis a realizar. Luego para cada una de las partículas se calcula la ubicación en la grilla construida. A partir de su posición se calcula la posición relativa al centro de dicha celda, esto se realiza creando una función para cada eje en donde si está en el centro de dicho eje, la función vale cero. Luego se calcula el porcentaje que le corresponde a la celda actual y a las celdas adyacentes, tal que si está en el centro de cierto eje, el peso en dicho eje es 1, mientras que si está en la unión de dos celdas, el peso en dicho eje es 0,5. Finalmente se asigna el porcentaje correspondiente a cada una de las celdas adyacentes, evitando revisar todas las posibilidades se hace uso de la función signo para saber en qué dirección moverse en cada eje.

### 4.3. Análisis y Resultados

Para los datos exportados de cada simulación se realizó una identificación de halos entre simulación, entre la de referencia (Planck) y las que tuvieron una variación del 5 %. Una vez realizada la identificación se procedió a hallar el cambio en velocidades y en cambio en las masas de cada halo a su “gemelo” en el universo correspondiente.

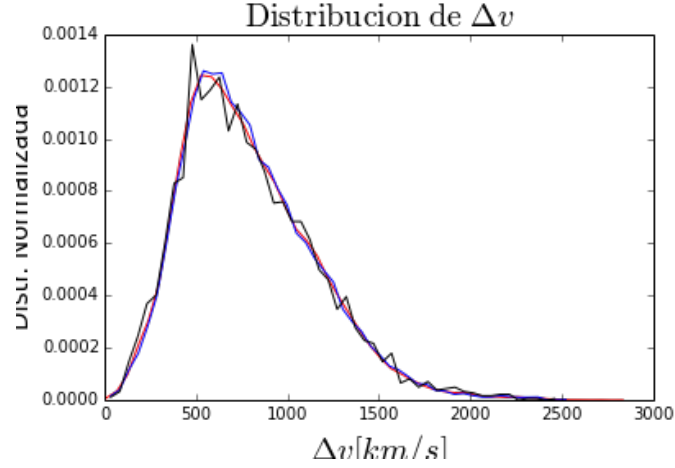


Figura 4.3: Distribución de cambios en la velocidad para  $\Omega - 5\%$

En las Figuras 4.4 y 4.3, se muestran las distribuciones normalizadas de las velocidades respecto al logaritmo de la masa en tres gráficas separadas por el rango de masas correspondientes. En estas gráficas se encontró un patrón inesperado y es que independientemente de la masa del halo, la distribución del cambio en velocidades es idéntica con un pico cercano a los  $500 km/s$ . Este dato es el más relevante puesto que cualquier experimento que quiera tener suficiente precisión y resolución para discernir entre cambios de 5 % en  $\Omega$  a través de redshifts, su método de medición de velocidades deberá ser capaz de discernir diferencias de velocidades en este orden.

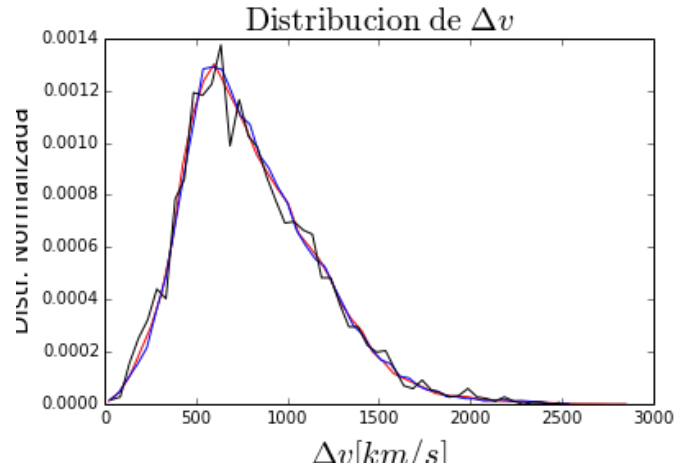


Figura 4.4: Distribución de cambios en la velocidad para  $\Omega + 5\%$

Para el caso de las Figuras 4.6 y 4.5 se puede ver que para los universos simulados la detección de los cambios en la densidad de materia en el universo de 5 % a través de mediciones en masas son



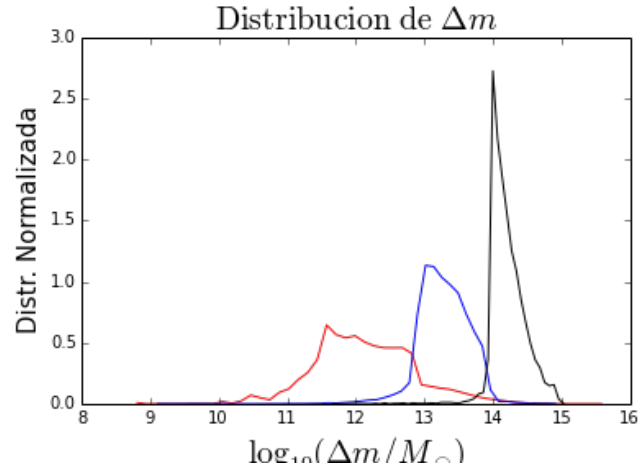


Figura 4.5: Distribución de cambios en la masa para  $\Omega - 5\%$

prácticamente inviables, puesto que los cambios en la masa de los halos varía en un factor cercano a 2. Un cambio en un factor de dos para métodos de medición de masas lejanas es casi indetectable, razón por la cual este método no sería el más recomendado.

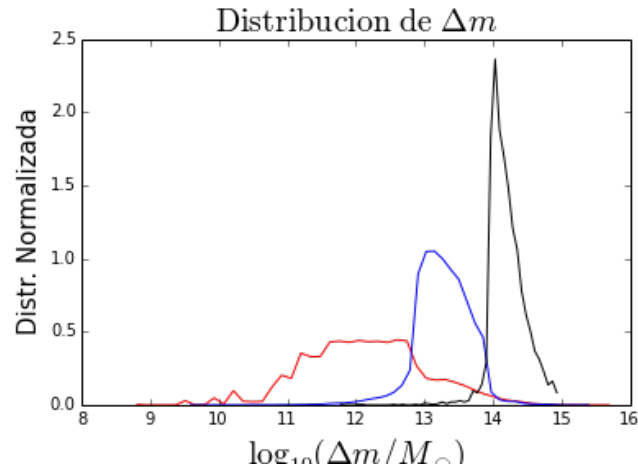


Figura 4.6: Distribución de cambios en la masa para  $\Omega + 5\%$

## 5. Conclusiones

En conclusión se observó que en caso de querer realizar análisis estadísticos para observaciones en el universo en pro de querer realizar medidas de los parámetros cosmológicos, una análisis mediante mediciones de redshift, los cambios que deberá detectar son del orden de  $\frac{\Delta v}{c} = 700/300000 \approx 2 \times 10^{-3}$ , algo que es alcanzable solamente con los experimentos de próxima generación, entre ellos DESI.

Además de esto, encontramos que un análisis mediante mediciones de masas sería inviable ya que los cambios de masa que caerían en variaciones de 5% en  $\Omega_0$  son cercanos a un factor de 2. Esto es indetectable con los métodos actuales, ya que los factores más bajos detectables con la tecnología disponible son cercanos a 10.

### 5.1. Trabajo Futuro

1. Tomar los resultados de las simulaciones resultado del proyecto para la realización de observaciones virtuales para cuantificar los efectos del *redshift space distortion* en observaciones.
2. Cuantificar en detalle la influencia de la red cósmica en la formación y alineación de halos de materia oscura con la estructura a gran escala.
3. Utilizar los resultados de las simulaciones como laboratorio de estudio de formación y evolución de poblaciones de galaxias observables con experimentos de siguiente generación, como DESI o HETDEX.

# Bibliografía

- [1] M. Baldi. Dark Energy simulations. *Physics of the Dark Universe*, 1:162–193, Nov. 2012.
- [2] S. Bhattacharya and A. Kosowsky. Dark energy constraints from galaxy cluster peculiar velocities. *Phys. Rev. D*, 77:083004, Apr 2008.
- [3] A. Font-Ribera, P. McDonald, N. Mostek, B. A. Reid, H.-J. Seo, and A. Slosar. Desi and other dark energy experiments in the era of neutrino mass measurements. *Journal of Cosmology and Astroparticle Physics*, 2014(05):023, 2014. 1
- [4] A. Guth. Was cosmic inflation the “bang” of the big bang? *The Beamline*, 1997,27. 3, 4
- [5] G. Hill, K. Gebhardt, E. Komatsu, N. Drory, P. MacQueen, et al. The Hobby-Eberly Telescope Dark Energy Experiment (HETDEX): Description and Early Pilot Survey Results. *ASP Conf.Ser.*, 399:115–118, 2008. 1
- [6] C. Knobel. An introduction into the theory of cosmological structure formation, 2012. 8
- [7] A. Loeb. *How did the first stars and galaxies form?* Princeton University Press, Princeton, NJ, 2010. 3, 4, 7
- [8] V. Springel. The cosmological simulation code gadget-2. *Monthly Notices of the Royal Astronomical Society*, 364, 2005. 9, 12