



***Facultad
de
Ciencias***

**Separación de componentes galácticas y
extragalácticas en astronomía de
microondas mediante técnicas de
combinación interna lineal**

**Separation of galactic and extragalactic
components in microwave astronomy using
internal lineal combination techniques**

**Trabajo de Fin de Grado
para acceder al**

GRADO EN FÍSICA

Autor: Sergio Blanco Gómez

Director: Diego Herranz Muñoz

Co-Director: R. Belén Barreiro

Septiembre - 2022

Abstract

In cosmology, the Cosmic Microwave Background plays a major role. Give its unique characteristics, it is key to comprehend crucial aspects about the history of the Universe, such as the formation of galaxies and cluster or estimating the cosmological parameters, needed to understand the evolution of the Universe. *Planck* mission has obtained, up to now, the best image of the Cosmic Microwave Background. To do so, complex diffuse source separation programs have been develop. They are a key piece of cosmology because it is an observational science. This means that every data collected is contaminated with not desired sources (foregrounds) that are present in the line of sight. This project intends to check the limits of a simple method, as ILC is, when applied to complex sets of data, like the *Planck* ones. The resultant map has a 10% error, which is not bad, but not enough to perform cosmology.

Resumen

En el campo de la cosmología, el Fondo Cósmico de Microondas juega un papel fundamental. Dadas sus peculiares características, un profundo conocimiento del mismo permite entender aspectos clave de la historia del Universo, como pueden ser la formación de estructuras como galaxias y cúmulos o la estimación de parámetros cosmológicos, cruciales para comprender cómo ha evolucionado el Universo. La misión Planck ha obtenido, hasta la fecha, la mejor imagen del Fondo Cósmico de Microondas. Para ello, se han desarrollado sofisticados métodos de separación de componentes difusas, cruciales en cosmología ya que, siendo una ciencia observacional, todos los datos están contaminados por *foregrounds*, fuentes que se encuentran en la línea de visión del observador, que son inevitables. Con este trabajo se quiere comprobar hasta qué punto un método simple como es ILC sirve aplicado a los complejos datos de *Planck*, resultando en un mapa con un error del 10%, que no está mal, pero para poder hacer cosmología no es suficiente.

Keywords: radioastronomía, resolución angular, Fondo Cósmico de Microondas, componentes difusas, separación de componentes, misión, *Planck*, fuentes puntuales, radioastronomy, angular resolution, Cosmic Microwave Background, diffuse components, component separation, *Planck* mission, point sources.

Índice

1	Introducción	2
1.1	Contexto histórico	2
1.1.1	Radioastronomía: origen	2
1.1.2	Diferencias de observación entre radio y visible	3
1.1.3	El fondo cósmico de microondas	4
1.2	El problema de la separación de componentes	6
1.2.1	Métodos de separación de componentes difusas	8
1.3	Modelo del cielo: <i>foregrounds</i>	9
1.4	Misión <i>Planck</i>	11
1.5	<i>Internal Linear Combination</i> : modelo matemático	13
1.6	Motivación y objetivos	15
2	Método experimental	15
2.1	Sistema operativo: máquina virtual	15
2.2	<i>Software</i> : Anaconda <i>python</i>	16
2.3	Datos utilizados	16
2.4	Acondicionamiento de los datos	18
2.5	Estructura del código	19
3	Resultados	20
3.1	Plantilla de separación	20
3.2	Sepración de componentes	21
3.3	Comparación de los resultados con el <i>CMB</i>	24
3.3.1	Mapa del residuo	25
3.3.2	<i>TT-plot</i>	28
3.4	Análisis de los resultados	29
3.5	Aplicación al mapa completo	29
4	Conclusiones	30
4.1	Valoración de los resultados	30
4.2	Opciones de mejora	31
4.3	Resultados del aprendizaje	32

1 Introducción

1.1 Contexto histórico

1.1.1 Radioastronomía: origen

En astronomía, el estudio de los cuerpos celestes se basa en todo el espectro electromagnético, lo que incluye una amplia región fuera del espectro visible (350 a 750 nm). Antes del siglo XX, dado que por motivos evolutivos el ser humano solo es capaz de percibir el mencionado espectro visible, toda la astronomía estaba basada en esta pequeña franja. Sin embargo, esto limita mucho el alcance de las observaciones, ya que el espectro de emisión de los cuerpos no está limitado únicamente al visible, requiriendo observarlos en todo el espectro para poder entenderlos en su totalidad. En muchos casos, la parte con la información más interesante esté fuera de esta región. Por ejemplo, la imagen de Sagitario A* (Figura 1.1), obtenida este mismo año, está tomada en radio [1], de forma que limitando la observación al visible, no se podría haber llegado a la misma.

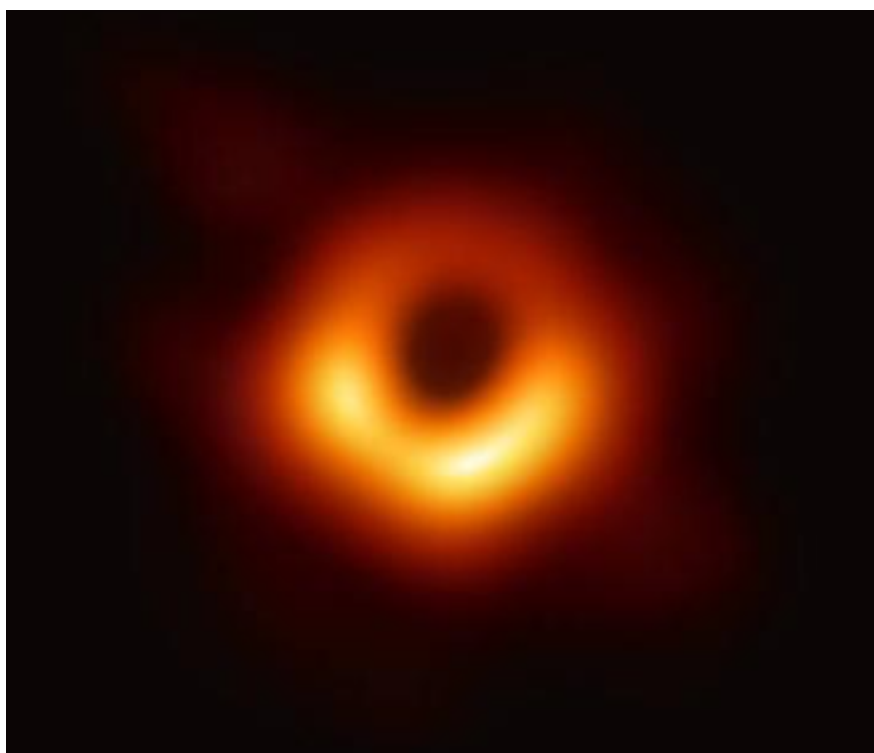


Figura 1.1: Imagen de Sagitario A*, obtenida en mayo de este año, por la colaboración *EHT* [2]

El campo de la radioastronomía surge en 1933 de manos de Karl G. Jansky que, estudiando fuentes de interferencia en comunicaciones para laboratorios Bell (que tenían problemas con sus sistemas de comunicación por motivos desconocidos), detectó una fuente de ruido extraña que interfería con su telescopio, el *merry-go-round* (Figura 1.2), concluyendo que provenía de la constelación de Sagitario en la Vía Láctea [3], región en la cual se encuentra Sagitario A*, el agujero negro supermasivo que se encuentra en el centro de la galaxia [4]. Esta fue la primera vez que se detectaron ondas de radio fuera del Sistema Solar.[5]

Sin embargo, la falta de profesionalidad científica de Jansky hizo que no fuese aún un campo establecido. Una de las claves para ello fue Grote Reber que, conociendo el descubrimiento de Jansky, construyó su propio radiotelescopio (Figura 1.3) realizando la primera detección sistemática de rayos cósmicos.

Gracias al desarrollo del radar durante la Segunda Guerra Mundial, en la década de los 60 se comenzaron a construir los primeros radio observatorios, destacando entre ellos el de la Universidad de Ohio, montado por John Kraus; o el conocido telescopio Arecibo, cuya construcción finalizó en 1963. Para 1964, ya era un campo establecido.

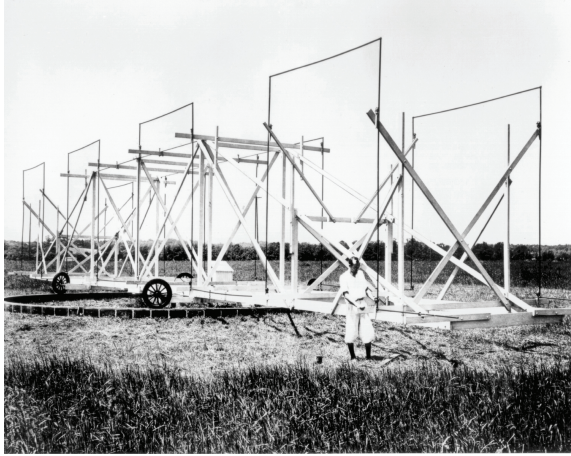


Figura 1.2: Fotografía del primer radiotelescopio, el *merry-go-round*, construido por Karl G. Jansky en 1933. [6]

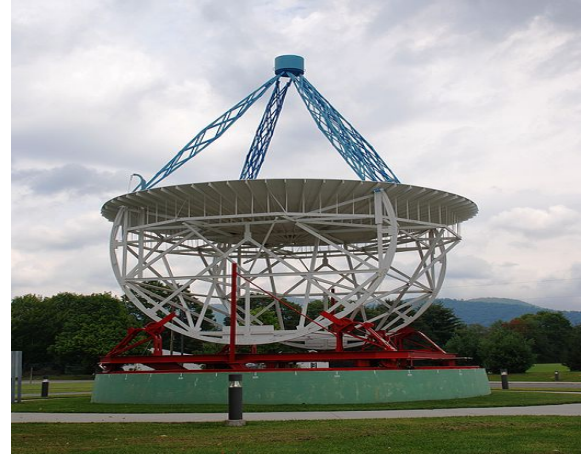


Figura 1.3: Fotografía del radiotelescopio construido por Grote Reber en 1937, con el que se realizaron las primeras detecciones sistemáticas de rayos cósmicos provenientes del interior de la Vía Láctea. [7]

1.1.2 Diferencias de observación entre radio y visible

A la hora de construir un telescopio, se debe tener en cuenta en qué frecuencias va a operar el mismo.

Uno de los parámetros más importantes a la hora de diseñar un instrumento de observación es la resolución angular, o capacidad de separar dos imágenes muy próximas, dada por [8]

$$\theta = 1,22 \frac{\lambda}{D} \quad (1)$$

Donde θ es la mencionada resolución angular, en radianes; λ es la longitud de onda de la fuente observada; y D es el diámetro, en este caso, del telescopio.

Sabiendo esto, se debe intentar minimizar esta magnitud para distinguir una mayor cantidad de cuerpos (cuanta más baja sea la resolución angular, más próximos podrán estar dos cuerpos distinguibles), por lo que, dado que las ondas de radio tienen una longitud de onda mucho mayor que la de la luz visible, se debe ampliar el tamaño del telescopio de radio para compensar la pérdida de resolución. Sin embargo, considerando el mejor de los casos, para igualar esta el segundo telescopio debe ser 40 veces más grande que el primero, lo que supone una gran inversión de materiales.

Por suerte, este problema se solventa gracias al fenómeno de interferencia, mediante el cual, usando varios telescopios similares para observar un mismo objeto, se puede resolver este usando el retardo de llegada a cada uno de ellos, consiguiendo así un "telescopio" con el tamaño del interferómetro (distancia que separa los telescopios). Es más, en casos extremos se ha llegado incluso a usar la trayectoria de la órbita terrestre como interferómetro, siendo de esta forma el diámetro

del "telescopio" máximo, si se trabaja desde la Tierra.

Comparando numéricamente ambos casos, al usar un interferómetro, como el *Very Large Telescope Interferometer (VLTI)*, se llegan a resoluciones (ecuación (1)) de 4 milisegundos de arco [9], mientras que al usar un único telescopio, como se hizo con los datos que se van a usar en este trabajo, recogidos por el telescopio *Planck*, son de varios minutos de arco [10]. Esta enorme diferencia refleja el gran poder que tiene el uso de la interferometría, así como qué escalas se pueden estudiar en caso contrario.

1.1.3 El fondo cósmico de microondas

Según la teoría de la relatividad, la velocidad máxima alcanzable en el universo es la velocidad de la luz, c . Para llegar a ella, un cuerpo masivo necesita de energía infinita, por lo que sólo partículas sin masa, como es el caso de los fotones (luz), como es lógico, viajan a dicha velocidad [11]. Aunque en un principio $3 \cdot 10^8$ m/s pueda parecer una velocidad muy elevada, realmente es un valor muy bajo a escalas cosmológicas, donde se trabaja con distancias de años-luz, incluso a primera escala (es decir, que los fotones, las partículas más rápidas del universo, tardan años en recorrer el espacio que separa a los cuerpos celestes). Es por ello que, en astronomía, observar lejos es equivalente a observar al pasado (distancia $\equiv$ tiempo).

Uno de los objetos de estudio más importantes en cosmología es el fondo cósmico de microondas (*CMB*, por sus siglas en inglés). Esto se debe a que es la fuente de radiación más antigua del universo. Más concretamente, como se observa en la Figura 1.4, se trata de fotones provenientes del desacoplamiento radiación-materia, es decir, cuando el universo se volvió transparente a la luz, tras 380000 años del *Big Bang*.

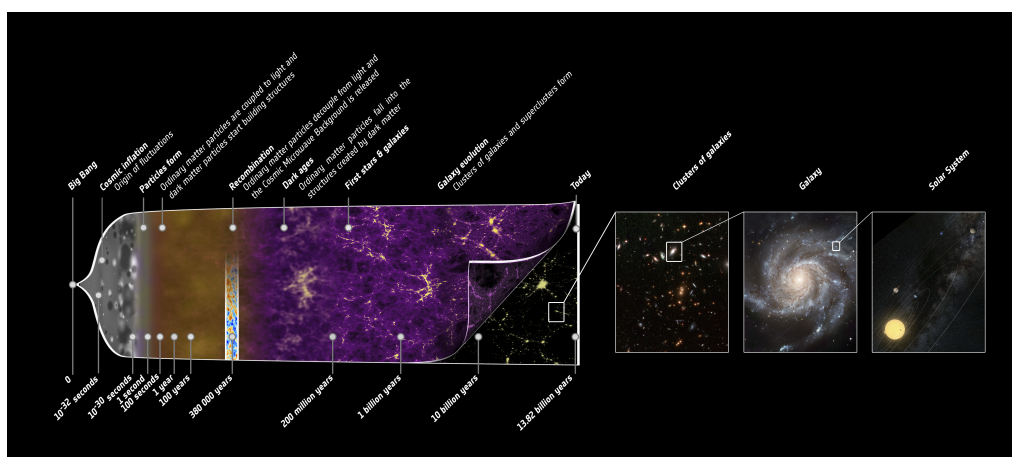


Figura 1.4: Esquema de la historia del universo, desde su origen con el *Big Bang* hasta la actualidad [12]

Las tres propiedades más importantes que presenta el *CMB* son homogeneidad, isotropía y emisión tipo cuerpo negro.

Comenzando por la tercera, que es la más sencilla de comprender, un cuerpo negro es aquel cuya absorptividad espectral es 1 (es decir, que no refleja nada de la luz incidente), por lo que su perfil de intensidad (Figura 1.5) sigue la ley de Planck [13]. Esto ocurre porque en la época previa al *CMB*, el Universo era tan denso y tenía una temperatura tan elevada que era opaco a la radiación (cualquier fotón se dispersaba a través de la interacción con una partícula dadas esas

condiciones, impidiendo su propagación [14]).

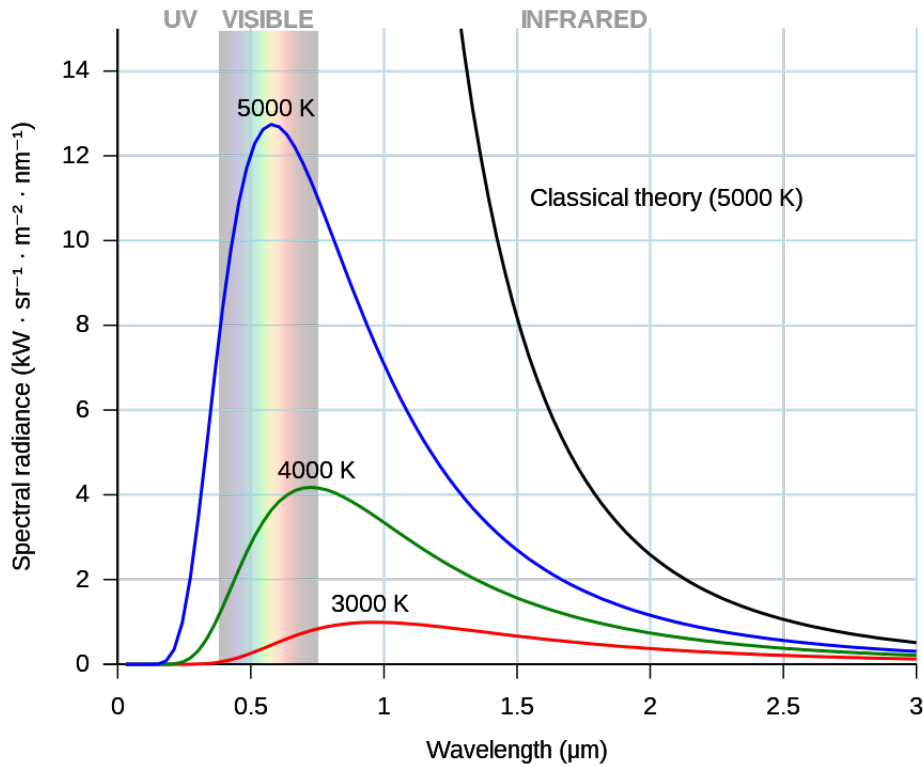


Figura 1.5: Perfil de intensidad de un cuerpo negro, en unidades de flujo espectral. [15]

Con respecto a la homogeneidad y la isotropía, son dos propiedades que fácilmente se confunden entre ellas, pero son diferentes e independientes la una de la otra. La homogeneidad significa que no hay puntos privilegiados, es decir, que todos los puntos presentan las mismas propiedades. Por su parte, la isotropía significa que no hay direcciones privilegiadas, por lo que no importa la dirección de observación, las propiedades van a ser las mismas [16]. Aunque en un principio pueda parecer que una cosa implica la otra, esto no es así, tal como ilustra la Figura 1.6.

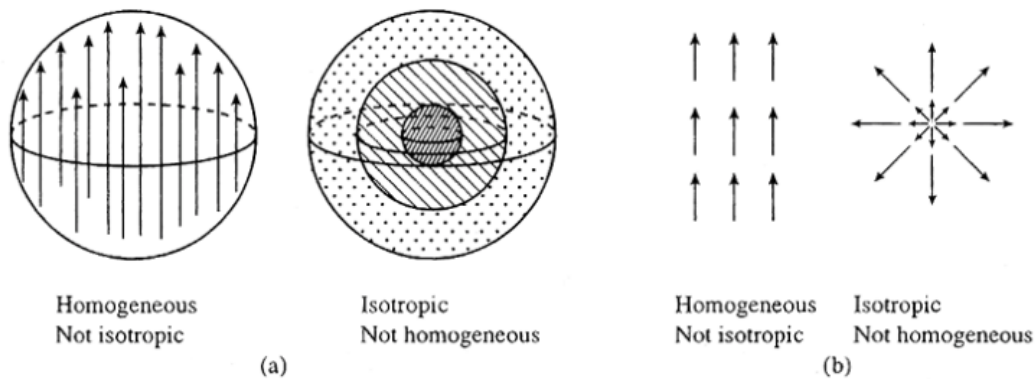


Figura 1.6: Casos en los que la homogeneidad y la isotropía no se cumplen simultáneamente. a) en 3D y b) en 2D [16]

Estas últimas propiedades, en el marco de la teoría del *Big Bang*, suponen lo que se conoce como el problema del horizonte. Para que el *CMB* sea homogéneo e isótropo, sus puntos deben estar conectados causalmente, cosa que es imposible ya que el tamaño del horizonte de partículas (distancia máxima que puede recorrer un fotón desde su emisión) en la época de la recombinación

(cuando se formó el *CMB*) es mucho menor que el radio del universo (es decir, que el universo es más grande que la máxima distancia que puede haber recorrido un fotón emitido cuando se formó el *CMB*), por lo que esta conexión es imposible.

El modelo inflacionario soluciona, además de otros problemas del modelo no relevantes en el contexto de este trabajo, el problema del horizonte. Partiendo de un universo a escala microscópica, donde, evidentemente, todos los puntos están conectados, este pasa por un periodo de expansión exponencial (se incrementa el tamaño en un factor de 10^{26} en apenas una fracción de segundo), antes de adquirir la velocidad de expansión predicha por la teoría del *Big Bang*. De esta forma, los puntos que previamente no podían estar conectados en realidad estaban mucho más cerca antes de la inflación, permitiendo esta conexión. [17]

La importancia del *CMB* en este campo radica en que, dado que esta radiación proviene de una época previa a la formación de estructuras como galaxias y cúmulos, midiendo las fluctuaciones en el mismo se puede comprender mejor cómo se formaron dichas estructuras. Además, siendo la fuente de radiación más antigua, contiene mucha información sobre las características del Universo, por lo que estas medidas permiten estimar los parámetros cosmológicos (curvatura del universo, velocidad de expansión, densidad de materia/materia oscura, etc).

El *CMB* fue descubierto por accidente en 1964 por dos radio astrónomos trabajando para laboratorios Bell, Arno Penzias y Robert Wilson (Figura 1.7), que, probando una antena de bocina para la detección de ondas de radio en el Proyecto Echo, se encontraron, habiendo eliminado todas las fuentes de interferencia conocidas (como radares y transmisiones de radio), con un exceso de radiación de origen desconocido. Tras descartar varias posibilidades, contactaron con un equipo de investigación de la Universidad de Princeton, que se encontraban trabajando en ese tema y creían que dicha radiación era el calor residual del *Big Bang*. Finalmente, en 1978 se les concedió el Premio Nobel a los dos astrónomos por su descubrimiento.

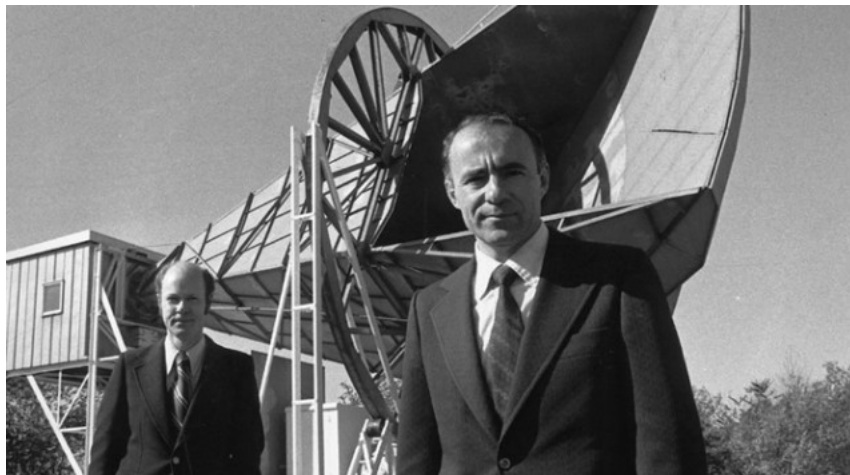


Figura 1.7: Fotografía de los astrónomos Arno Allan Penzias (derecha) y Robert Woodrow Wilson (izquierda), junto a la antena con la que detectaron por primera vez el *CMB* [18]

1.2 El problema de la separación de componentes

A diferencia de lo que sucede en otras disciplinas, cuando se realiza una observación en astronomía no es posible modificar la disposición de los cuerpos a observar, ni cambiar el ángulo de observación para sortear posibles elementos que se encuentran en primer plano (llamados *foregrounds*), tal como se observa en la Figura 1.8. Es por ello que los datos recogidos no corresponden única-

mente a aquel cuerpo que se quiere estudiar, sino que estos *foregrounds* se encuentran mezclados en la misma imagen. La Figura 1.9 ilustra este problema con un ejemplo muy visual, en el que dos galaxias que se encuentran en la misma línea de visión aparecen solapadas en la imagen, aunque realmente son dos galaxias independientes.

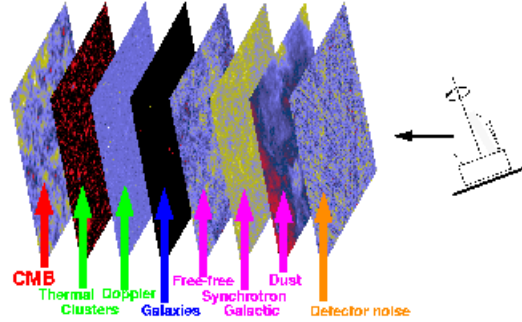


Figura 1.8: Representación esquemática del problema de los *foregrounds*. El observador, situado a la derecha, está interesado en obtener una imagen del *CMB*, pero hay otras fuentes en el camino como galaxias, polvo, el propio ruido del detector, etc. [19]



Figura 1.9: Imagen de dos galaxias solapadas (*NGC 3314*), fruto de un efecto visual al encontrarse en la misma línea de visión, una delante de la otra. [20]

En principio, este problema no se puede resolver. Se parte de una serie de mapas, y_i , mediante los cuales se quiere recuperar los mapas de emisión de las fuentes presentes en él, s_i . La ecuación que describe estos mapas de observaciones es

$$y(p) = As(p) + n(p) \quad (2)$$

Donde A es la previamente mencionada matriz de mezcla (que relaciona las fuentes entre sí), y n es el ruido. De esta forma, la separación de componentes consiste en invertir la ecuación (2).

Lo mas sencillo es simplemente invertir la matriz A , obteniendo una estimación de los componentes

$$\hat{s} = A^{-1}y = s + A^{-1}n \quad (3)$$

pero, debido a la presencia del ruido, no siempre es la mejor aproximación al problema. Además, para poder hacer esto se debe conocer la matriz de mezcla, para lo cual es requisito imprescindible conocer cuáles son y cómo se comportan las fuentes contaminantes, así como otros problemas que se discutirán en 4.2.

Existen otras formas de realizar esta misma inversión, en función del operador W usado para despejar s (en el caso anterior, que era el más simple, $W = A^{-1}$), que requieren de diferentes condiciones de los datos (número de datos/mapas, conocimiento del ruido n , etc.), todos ellos tratando de minimizar el error de $\hat{s}$. Entre ellos destacar el filtro Wiener. A diferencia de otros métodos, la solución de este no toma la forma $\hat{s} = s + Wn$, sino que hay un término multiplicando a la matriz de fuentes s , por lo que puede haber residuos de otros foregrounds en el resultado final. Es decir, el mapa que minimiza la varianza del resultado no necesariamente es aquel en el que todo el residuo viene causado por el ruido.

Para terminar, el perfil de intensidad $I(\nu)$ de la mayoría de las fuentes es bien conocido, comportándose siempre en base a la misma función. Sin embargo, las galaxias no pueden ser descritas tan fácilmente, ya que, como cada una presenta una proporción distinta de fuentes emisoras, su perfil es único, y no es posible generalizar su descripción como en el resto de casos. Es por ello que, a la hora de aplicar este tipo de métodos, se suele primero eliminar de forma independiente las diversas galaxias de la imagen, para posteriormente poder aplicarlo al resto de componentes, sin que estas afecten en gran medida al resultado.

1.2.1 Métodos de separación de componentes difusas

Todos los métodos de un punto común, que es la descripción de los datos recogidos como

$$d_\nu = b_\nu * x_\nu + n_\nu \quad (4)$$

donde d_ν , x_ν y n_ν son las observaciones, emisiones y ruido a frecuencia ν , respectivamente, mientras que b_ν es el rayo (asumiendo que es gaussiano) a la misma frecuencia. El problema radica en el término x_ν , que es una superposición de componentes, generalmente asumida como una combinación lineal definida a través de la matriz de mezcla A (se hablará más en detalle de ello en 4.2).

Con esto aclarado, estos son algunos de los métodos utilizados para la separación [10]:

-Muestreo de Gibbs: método usado por Commander, uno de los códigos oficiales de *Planck*, se basa en ajustar directamente un modelo de los componentes de la imagen (*CMB*, *foregrounds* y ruido) a la temperatura de la antena en mapas de baja resolución.

-Análisis de componentes correlacionadas: se realiza una estimación de la matriz de mezcla A a través de correlaciones espaciales de los datos, así como restricciones de plantillas externas y modelos de los *foregrounds*. Con ello, se reconstruyen los componentes usando un filtro Wiener.

-Análisis de componentes independientes: Sin realizar suposiciones previas, salvo que los componentes son independientes, se obtiene la matriz de mezcla A maximizando la no-gaussianidad en cada punto.

-**Método de máxima entropía en el espacio armónico:** asumiendo una distribución de probabilidad de máxima entropía dados los componentes de la imagen, se estiman los mapas ajustando unos pesos con modelos y espectros de los *foregrounds*.

-**SEVEM *Spectral estimation via expectation maximisation*:** primero, se reducen los *foregrounds* con una resta interna. Con estos mapas con *foregrounds* reducidos se estima el espectro del *CMB*, para finalmente aplicar un filtro Wiener sobre los mismos.

Y un largo etcétera. En resumen, hay multitud de formas de abarcar este problema, partiendo de más o menos asunciones, optando por una u otra en función de qué se prioriza obtener y las características del proyecto.

En el caso de este trabajo, se ha optado por usar *ILC* porque es un método sencillo, que no requiere de una programación muy compleja ni un ordenador muy potente. Además, conceptualmente también es muy sencillo, y no requiere de prácticamente nada de información previa. Evidentemente esto supone ciertos problemas. El más importante a mencionar es que, al no usar la matriz de mezcla A , no se tienen en cuenta posibles cambios en distintas zonas del mapa, que son una realidad. Dadas la resolución de los datos de *Planck*, mayor que misiones previas, este tipo de problemas se acentúan, requiriendo de métodos más sofisticados como los previamente descritos.

1.3 Modelo del cielo: *foregrounds*

Antes de empezar a realizar el trabajo en sí, es conveniente describir cuáles son los diferentes *foregrounds* que van a estar presentes en la imagen de trabajo, al menos, los más importantes [21]. La Figura 1.10 muestra cuáles son, así como su perfil de temperatura.

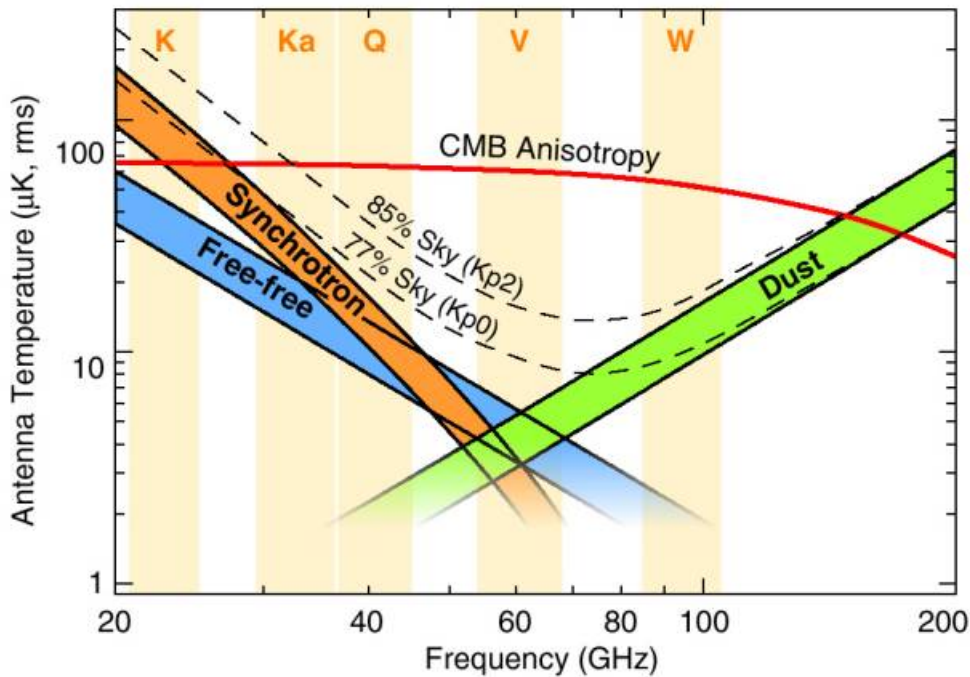


Figura 1.10: Diagrama de la dependencia de los *foregrounds* más importantes a tener en cuenta en el trabajo [22]

En primer lugar, la emisión sincrotrón la causan partículas cargadas moviéndose en un cam-

po magnético, en el que siguen trayectorias curvas (concretamente espiral). Dado que se tiene una partícula cargada acelerada, se emite radiación. En la Vía Láctea, los campos magnéticos responsables de este efecto se extienden más allá del plano galáctico y, dado que los electrones más energéticos pueden escapar de dicha región, se observa emisión sincrotrón incluso a altas latitudes, tal como se observa en la Figura 1.11.

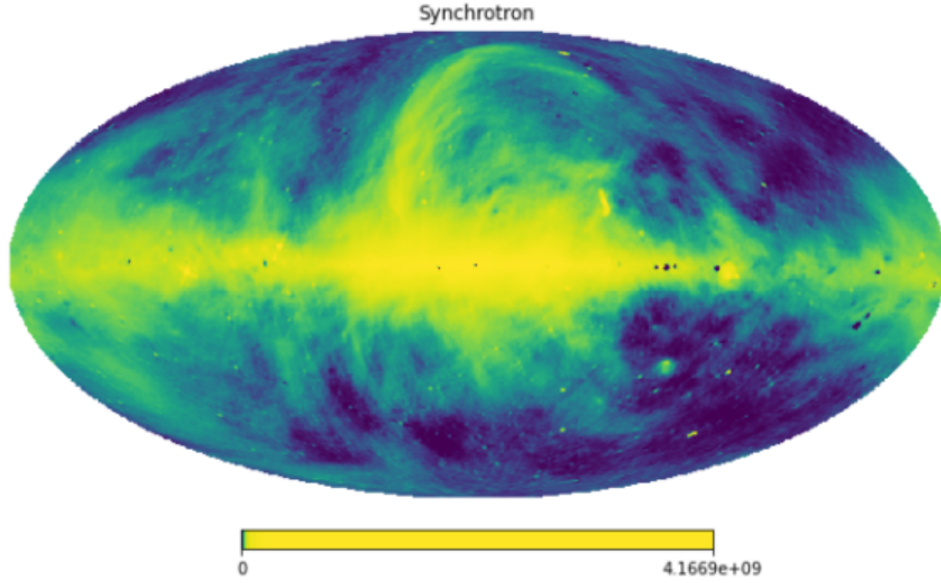


Figura 1.11: Mapa de la radiación sincrotrón, obtenido por Commander [23], uno de los cuatro códigos de separación de componentes oficiales de la colaboración *Planck*.

La emisión libre-libre, la mas tenue de las tres consideradas, está causada por la interacción a través de efecto *Bremsstrahlung* de los electrones libres (de ahí su nombre) con iones pesados. Este efecto consiste en un frenado de los electrones debido a estos iones, emitiendo radiación en el proceso, y se origina en el medio interestelar, principalmente en rangos de longitud de onda en torno al centímetro. La Figura 1.12 muestra el mapa de esta emisión.

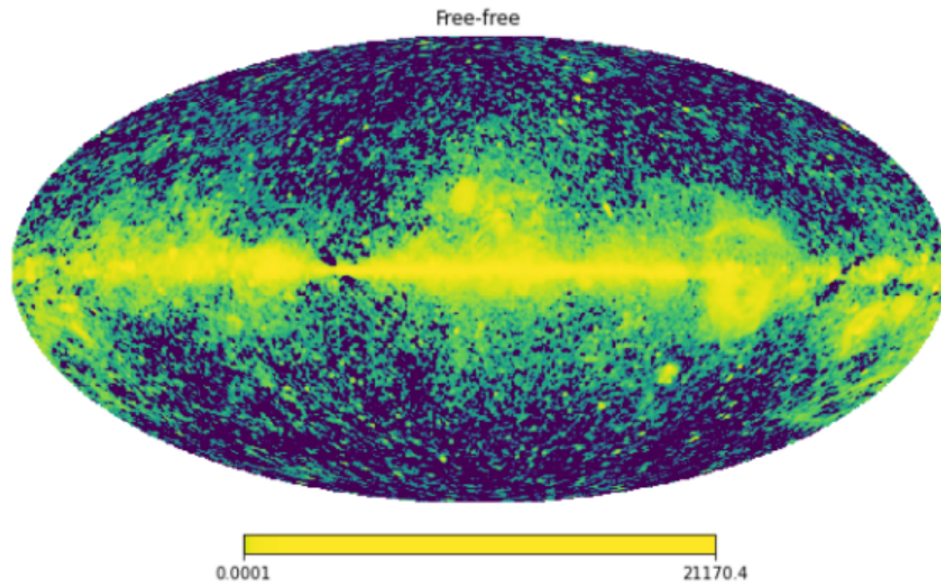


Figura 1.12: Mapa de la radiación libre-libre, obtenido por Commander [23], uno de los cuatro códigos de separación de componentes oficiales de la colaboración *Planck*.

Por último, se debe tener en cuenta el polvo, presente en el espacio interestelar, que emite principalmente en radio/infrarrojo. Este se compone de pequeñas partículas de distintos tamaños

y formas, pudiendo medir desde unos pocos nanómetros hasta micrómetros. Aunque puede emitir a través de varios procesos físicos, en el caso del estudio del *CMB* el más importante es la emisión de cuerpo gris en el infrarrojo lejano (un cuerpo gris es aquel que, emitiendo según la ley de radiación del cuerpo negro, no tiene absorptividad igual a 1, es decir, que sigue el mismo perfil pero con una intensidad menor de la que tendría un cuerpo negro [24]). También se puede detectar por efectos de absorción en la región ultravioleta, pero no es importante en este caso, ya que no se trabaja en esas frecuencias. En la Figura 1.13 se muestra el mapa de la emisión debido al polvo.

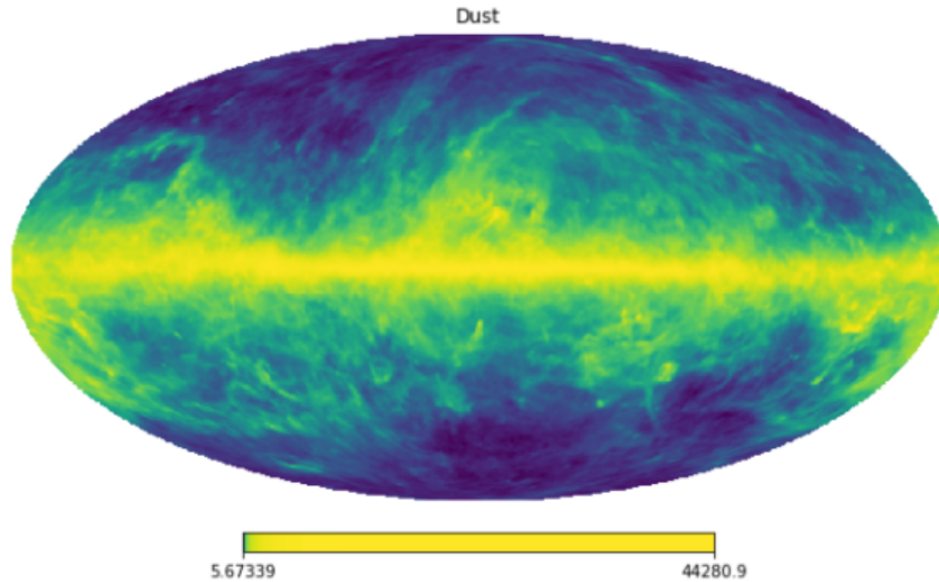


Figura 1.13: Mapa de la radiación del polvo, obtenido por Commander [23], uno de los cuatro códigos de separación de componentes oficiales de la colaboración *Planck*.

Observando estos mapas, se observa claramente como hay ciertas regiones más contaminadas que otras (destacando sobre todo la zona del plano galáctico, la franja horizontal central), pudiendo perjudicar al resultado de otras zonas. Por ello, una solución es dividir el mapa del cielo en regiones, de forma que se tratan de forma independiente las regiones con más y menos contaminación, dando lugar a mejores resultados. El problema de esto es el problema de la discontinuidad. Al realizar esta división, en las zonas del mapa fronterizas entre las regiones en las que se ha dividido, los resultados no van a ser continuos en una zona y en otra.

1.4 Misión *Planck*

Uno de los proyectos más importantes en el campo de la cosmología ha sido la misión *Planck*, destinada al estudio del fondo cósmico de microondas, cuya importancia fue descrita en 1.1.3.

Lanzado en mayo de 2009, y colocado en órbita, al igual que otras misiones de la *ESA* (como el *Webb*, actualmente operativo), en el segundo punto de Lagrange (L2) (Figura 1.15), el telescopio *Planck*, de 4,2m de diámetro 1.14 era, hasta la fecha, el instrumento destinado a la observación *CMB* con mayor sensibilidad, rango de frecuencias y resolución angular, hasta tal punto que su precisión no estaba limitada por el aparato en sí, sino por los propios límites astrofísicos. Esto que explica que, a día de hoy, los resultados obtenidos tengan todavía un gran valor.

Este telescopio consta de dos instrumentos que observan en 9 canales de frecuencia: *LFI* (*Low Frequency Instrument*), entre 30, y 70 GHz; y *HFI* (*High Frequency Instrument*), entre 100 y 857 GHz. Los detectores son, en el primer caso radiómetros (usa electrones de alta movilidad, detectándose en transistores); y en el segundo, bolómetros (detectores dopados con germanio que

realizan transformaciones de neutrones) [25]. Las resoluciones y niveles de ruido de cada canal, así como las frecuencias en las que trabaja cada aparato, se muestran en la Tabla

Tabla 1.1: Valores de resolución angular y ruido en cada canal de los observados por el telescopio *Planck* [26]. Por encima de la división, frecuencias de observación de LFI, y por debajo, las de *HFI*

ν (GHz)	FWHM(arcmin)	ruido(μK_{CMB} deg)	ruido(KJy/sr deg)
30	32.29	2.5	
44	27.94	2.7	
70	13.08	3.5	
100	9.66	1.29	
143	7.22	0.55	
217	4.90	0.78	
353	4.92	2.56	
545	4.67		0.78
857	4.22		0.72

La misión llegó a su fin el 23 de octubre de 2013, lanzándose los resultados finales en 2018, consiguiendo gracias a ello, entre muchas otros logros, el cálculo más preciso hasta entonces de la edad del universo.[27] [28]

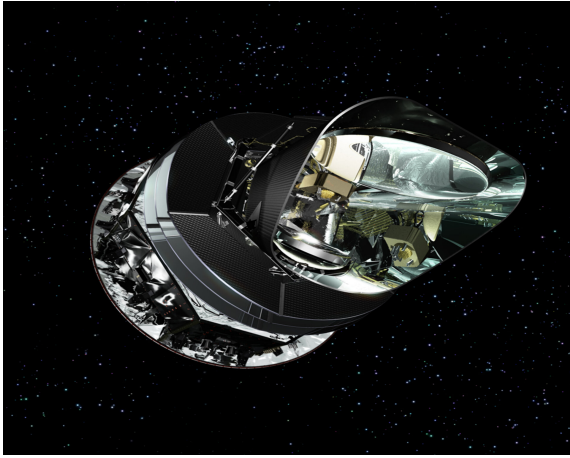


Figura 1.14: Telescopio *Planck*, usado para la recolección de datos de la misión con el mismo nombre [29]

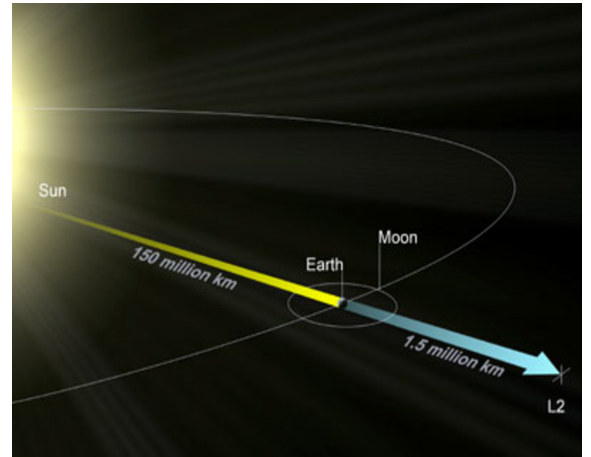


Figura 1.15: Esquema explicativo de la localización del segundo punto de Lagrange (L2)

Cabe destacar el trabajo del IFCA en esta misión [30], participando en el desarrollo del instrumento de baja frecuencia, *LFI*, así como desarrollando uno de los cuatro códigos oficiales de la misión, SEVEM, cuyo resultado final se muestra en la Figura 1.16.

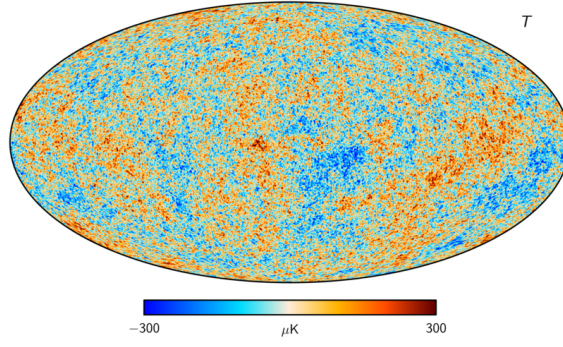


Figura 1.16: Mapa de temperatura del *CMB* reconstruido a partir de los datos de *Planck* con SEVEM, el código desarrollado por el equipo de trabajo del IFCA [31]

1.5 *Internal Linear Combination: modelo matemático*

Dadas las limitaciones en este proyecto con respecto al realizado por los equipos que trabajaron en la misión *Planck*, se ha utilizado un modelo mucho más sencillo.

El punto de partida es un mapa de datos d , en el que se encuentran mezclados la señal de interés s , esto es, el componente que se quiere extraer del resto del mapa (en este caso concreto, el *CMB*); un conjunto de *foregrounds* f , que son aquellas fuentes que se encuentran en la línea de observación, y la principal causa de la importancia de estos métodos; y ruido n . De esta forma, asumiendo que todas las fuentes son independientes entre sí, la ecuación que describe el problema es:

$$d_i(p) = s_i(p) + f_i(p) + n_i(p) \quad (5)$$

Donde el subíndice i hace referencia al canal o mapa (frecuencia de detección); y la variable p , al píxel.

De esta forma, el punto de partida son los datos deseados con un 'error' (entendido como señales no deseadas ¹) proveniente del ruido y otras fuentes de emisión (el previamente mencionado *foreground*).

El objetivo del método es, lógicamente, minimizar este error sin modificar la señal deseada (es decir, sin tocar la señal del *CMB*, deshacerse de los *foregrounds* y los ruidos). Dado que el *CMB* es constante en el dominio de la frecuencia, una forma de llegar a dicho resultado es hacer una media ponderada, $\hat{s}$, que minimiza ese 'error' sin modificar la señal del *CMB*. En realidad esto no es el mapa del *CMB* como tal, sino la estimación del mismo que más se aproxima a la señal original, s .

$$\hat{s}(p) = \sum_i^{N_\nu} w_i d_i(p) \quad (6)$$

siendo w_i el peso asociado al canal i , que deben cumplir

$$\sum_i w_i = 1 \quad (7)$$

y se obtienen a partir de la matriz de covarianza C , definida como

¹En este trabajo sólo se pretende recuperar la imagen del *CMB*, dando igual lo que ocurra con los *foregrounds*. Usando métodos más sofisticados que permitan recuperar también estos últimos, en ningún caso podrían ser considerados como un 'error'.

$$C_{ij} = \langle d_i d_j \rangle = \frac{1}{N_{pix}} \sum_{p=1}^{N_{pix}} (d_i(p) - \overline{d(p)})(d_j(p) - \overline{d(p)}) \quad (8)$$

donde $\overline{d(p)}$ es la media de todos los canales en el píxel p .

Aplicando esto a la ecuación (5), se llega a que la estimación es

$$\hat{s} = s + \sum_i w_i f_i + \sum_i w_i n_i \quad (9)$$

Con estos operadores definidos, para llegar a $\hat{s}$ se debe minimizar la varianza, dada por

$$Var(T) = \langle T^2 \rangle - \langle T \rangle^2 = w^T C w \quad (10)$$

De esta forma el objetivo es minimizar la función $f(w)$, definida a continuación, bajo la restricción $g(w)$ (ecuación (7)).

$$f(w) = \sum_{i,j=1}^k w_i C_{ij} w_j \quad (11)$$

Para ello, según el método de multiplicadores de Lagrange [32], se debe encontrar el punto w_0 que satisface

$$\nabla f(w_0) = \lambda \nabla g(w_0) \quad (12)$$

siendo λ una constante arbitraria.

Las derivadas de f se calculan a partir de la ecuación (11), mientras que las de g son 1.

$$\frac{\partial f}{\partial w_i} = 2 \sum_{j=1}^k C_{ij} w_j \quad (13)$$

El mínimo de f se encuentra resolviendo el sistema de ecuaciones definido en (12), al tiempo que se cumple la condición de la ecuación (7). En forma matricial, esto es resolver

$$\begin{bmatrix} 2\mathbf{C} & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} \quad (14)$$

Donde la negrita se ha usado para matrices/vectores, distinguiéndolos de los escalares.

Resolviendo la ecuación (14), se llega a que los pesos vienen dados por

$$w_i = \frac{\sum_j C_{ij}^{-1}}{\sum_{i,j} C_{ij}^{-1}} \quad (15)$$

y, con estos, aplicando (6), se obtiene la estimación del mapa deseada.

1.6 Motivación y objetivos

Al igual que la propia misión *Planck* en la que se basa este trabajo, la principal motivación para llevarlo a cabo es la mejor comprensión de la historia de nuestro universo a través de uno de los elementos de mayor importancia en su estudio, el fondo cósmico de microondas, así como experimentar de primera mano como se trabaja con los datos brutos en este campo. Como se comentó en 1.4, las características del instrumento utilizado, además del hecho de que se dedicó íntegramente al *CMB*, hacen que estos datos sean de los mejores para este trabajo, si no los que más.

Como es lógico, no se espera obtener unos resultados tan precisos como los de un gran equipo trabajando durante años en cosa de dos meses, sino más bien la experiencia de trabajar con estos datos, así como determinar hasta que punto puede ser más eficiente el uso de métodos simples, frente a los necesarios cuando se busca máxima precisión, para proyectos de menor calibre, en los que resultados no tan precisos ejecutados en una menor cantidad de tiempo es preferible, tratando de proponer añadidos al método base que mejoren la calidad de los resultados sin sacrificar tiempo. A título personal, este paso sirve para experimentar en su forma más básica un trabajo real (hablando incluso fuera del ámbito científico), en el que, a diferencia de los problemas académicos con una única solución, accesible mediante un método concreto, previamente conocido, los problemas que surgen se pueden afrontar de infinitud de formas, todas ellas perfectamente válidas si conducen a un resultado satisfactorio.

Como se comentó en 1.2.1, las características de los datos de *Planck* hacen que el resultado sea más sensible al método utilizado que en proyectos anteriores, como WMAP (*Wilkinson Microwave Anisotropy Probe*) [33], en el que se usó precisamente *ILC*. Por ello, se quiere probar hasta qué punto es válido este método, además de como afecta el usarlo sobre datos más complejos.

Con respecto a la cosmología y la astrofísica, se pretende, como ya se ha mencionado anteriormente, comprender mejor cómo funciona el universo, teniendo en mente el objetivo final de conseguir un mapa del *CMB*, al tiempo que se observan diversas simulaciones, teniendo así una mejor comprensión de las diversas componentes que forman los datos brutos recogidos en una misión espacial, y como afecta cada una de ellas a la imagen inicial.

En cuanto a la programación, a pesar de tratarse de un código bastante sencillo, sigue siendo, por varios órdenes de magnitud, más complejo que los realizados en un experimento al uso durante la carrera (en la mayoría de los casos, estos últimos consistían simplemente en introducir una fórmula y hacer un ajuste sencillo). Por ello, este trabajo ayuda a mejorar a la hora de elaborar un programa, siendo la primera experiencia con datos reales de un gran proyecto que requieren de métodos específicos de almacenamiento, añadiéndole a ello una preocupación real por la optimización debido al volumen de datos, además de comprender la importancia de la documentación, no solo para otra persona que vaya a utilizar el programa, sino para uno mismo, en caso de que haya necesidad de reutilizar partes del código, así como para facilitar el tratamiento de errores.

2 Método experimental

2.1 Sistema operativo: máquina virtual

Como se comentará más adelante, algunas de las librerías utilizadas para el programa no son compatibles con el sistema operativo Windows. Por ello, todo el desarrollo del código se ha llevado a cabo en una máquina virtual, concretamente 'Oracle VM VirtualBox', usando el sistema

operativo 'Ubuntu 22.04'.

La configuración usada en la máquina ha sido de 8GB de RAM, de forma que queden recursos disponibles para el sistema fuera de ella con el fin de realizar consultas, por motivos de comodidad; y 500 GB de almacenamiento, necesarios para almacenar los datos brutos, de varias decenas de GB de peso.

2.2 Software: Anaconda *python*

Python es un lenguaje de programación muy utilizado en el ámbito científico, que cuenta con muchas librerías destinadas al análisis de datos en multitud de ramas. Esto, sumado a una licencia de código abierto (que facilita la distribución de los programas creados, evitando barreras económicas), así como una gran simplicidad de programación, no por ello perdiendo eficiencia (en parte debido a que muchas de las librerías, como *numpy*, están programadas en C, más rápido), hacen que *python* sea una de las mejores opciones a utilizar.

Para realizar la instalación, Anaconda [34] es la opción más sencilla, ya que proporciona el intérprete *Spyder*, así como una serie de librerías de uso habitual, como pueden ser *numpy* o *matplotlib*, entre otras muchas.

En este trabajado concreto, las dos librerías fundamentales en el desarrollo del código son *astropy* y *Healpy*. La primera, como su propio nombre indica, está diseñada para trabajar en el campo de la astrofísica, y es muy útil para manejar los archivos FITS, especialmente la información de sus encabezados. Por su parte, la segunda es específica para datos en formato HEALPix (desarrollado en 2.3), permitiendo representar mapas como el de la Figura 1.16, así como haciendo transformaciones de coordenadas.

2.3 Datos utilizados

Los datos utilizados se encuentran disponibles en la página web de *Planck Legacy Archive (PLA)* [23], tanto los mapas resultantes de su trabajo, usados como modelo para verificar la validez del método, así como las simulaciones sobre las cuales se aplica el mismo.

Estos datos son ficheros FITS (*Flexible Image Transport System*) [35], un formato ampliamente usado en el campo de la astrofísica, formados por un encabezado de *keywords* en formato ASCII (legible para personas), que describe los datos que se encuentran a continuación (formato de ordenación y tamaño del mapa, magnitud y unidades de los datos, componentes incluidos...), encriptados, lo que permite el almacenamiento de grandes cantidades de datos de manera eficiente.

En algunas misiones, el formato de pixelización usado en los mapas es HEALPix [36], consistente en dividir la esfera celeste en 12 regiones cuadradas iguales, tal como se muestra en la Figura 2.17, caracterizadas por el número de píxeles en cada lado de este cuadrado, N_{side} , que es una potencia de 2.

Esto es así porque se está observando el cielo completo, una superficie esférica. Cuando se tiene una imagen, los píxeles deben cubrirla de manera uniforme y regular (es decir, que todos los píxeles sean iguales y sus dimensiones encajen con las de la imagen). En una imagen plana esto se consigue fácilmente usando cuadrados, pero al trabajar con una esfera se debe usar otra geometría. HEALPix es el estándar en el *CMB* por las propiedades matemáticas de su simetría,

que permiten realizar fácilmente transformaciones de armónicos esféricos.

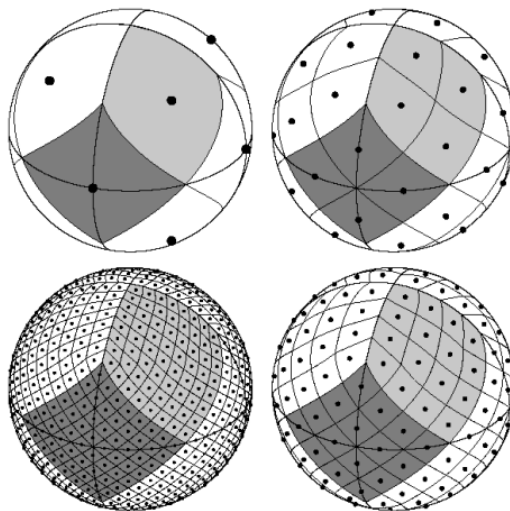


Figura 2.17: Vista de la forma de dividir la esfera de HEALPix para varios valores de lado. En sentido horario, partiendo de la imagen superior izquierda, este vale $N_{side} = 1, 2, 4, 8$, lo que equivale a un número de píxeles $N_{pix} = 12 \cdot N_{side}^2 = 12, 48, 192, 768$, respectivamente. [36]

En el caso de la misión *Planck*, los mapas son fluctuaciones, bien de temperatura (K), bien de densidad de flujo espectral (Jy/sr)², ya que, dependiendo de la frecuencia de observación, debido al método empleado para la recogida de datos es mejor usar una magnitud u otra [25]. En la tabla 2.2 se muestra un ejemplo de los principales parámetros en la simulación parcial (incluyendo solo algunos componentes) a 100GHz.

Con respecto a las unidades de trabajo, los datos son mapas de fluctuaciones (es decir, que mide la variación con respecto al valor medio, como se puede observar en la Figura 1.16), en frecuencias que van desde 30GHz hasta 857GHz y las unidades son bien de temperatura (K), bien de densidad de flujo espectral (Jy/sr), dependiendo de la frecuencia de trabajo, ya que, dependiendo de la misma, el método de observación hace que sea más sencillo trabajar con unas unidades o con otras. Además, como se ve en la Tabla 1.1, la resolución no es igual en todos los canales, y también hay que tratar los datos en este aspecto para igualarlas.

²1Jy = $10^{-26} \text{W} \cdot \text{m}^{-2} \cdot \text{Hz}^{-1}$

Tabla 2.2: *Keywords* más importantes en una de las simulaciones utilizadas, que sirven para definir los datos presentes en un archivo FITS. [37][38]

<i>Keyword</i>	Valor	Descripción	<i>Keyword</i>	Valor	Descripción
BITPIX	8	Número de bits por cada dato	TUNIT3	K_CMB	Unidades de la columna 3
NAXIS	2	Número de ejes	POLCCONV	COSMO	Convenio de polarización de HEALPix
NAXIS1	12288	Número de bytes en el eje 1	COORDSYS	G	Sistema de coordenadas
NAXIS2	12288	Número de bytes en el eje 2	PIXTYPE	HEALPIX	Tipo de pixelización
TFIELDS	3	Número de columnas	ORDERING	RING	Formato de ordenación
TTYPE1	TEMPER...	Número de la columna 1	OBJECT	FULLSKY	Todo el cielo o parcial
TUNIT1	K_CMB	Unidades de la columna 1	FIRSTPIX	0	Índice del primer píxel
TTYPE2	Q_POLAR...	Número de la columna 2	LASTPIX	12582911	Índice del último píxel
TUNIT2	K_CMB	Unidades de la columna 2	NSIDE	1024	Número de píxeles por lado
TTYPE3	U_POLAR...	Número de la columna 3	NPIX	12582912	Número de píxeles
			BEAMTYPE	GAUSSIAN	Tipo de haz observado

2.4 Acondicionamiento de los datos

Para realizar el cambio entre unidades, se debe aplicar la ley de Planck de la radiación del cuerpo negro [13]

$$B(\nu) = \frac{2h\nu^3}{c^2} \frac{1}{\exp(\frac{h\nu}{k_B T}) - 1} \quad (16)$$

Dado que los mapas son de fluctuaciones, se debe realizar la transformación aplicando derivadas

$$dB = dT \left(\frac{dB}{dT} \right) |_{T_0} \quad (17)$$

Donde T_0 es la temperatura media del mapa, que en este caso es la temperatura del *CMB*, $T_0 = 2,7255\text{K}$ [39].

De esta forma, la ecuación que describe la transformación de unidades viene de combinar las ecuaciones (16) y (17)

$$\Delta T = \Delta B \frac{(\exp(\frac{h\nu}{k_B T_0}) - 1)^2}{\exp(\frac{h\nu}{k_B T_0})} \cdot \frac{c^2 k_B T_0^2}{2h^2 \nu^4} \quad (18)$$

Por otro lado, las resoluciones angulares de los distintos canales, como se ve en la Tabla 1.1, no es la misma. Esto hace que no se puedan sumar los datos sin más, ya que se obtendrían resultados incoherentes, sino que hay que hacer un proceso de suavizado antes. Este consiste en, dada la resolución en cierto canal, aumentar su anchura a media altura (*FWHM*, Full Width at Half Maximum) de forma que, al final, todos tengan la misma.

2.5 Estructura del código

La forma en que se ha estructurado el programa se describe en el diagrama de secuencia (*sequence chart* en inglés) de la Figura 2.18.

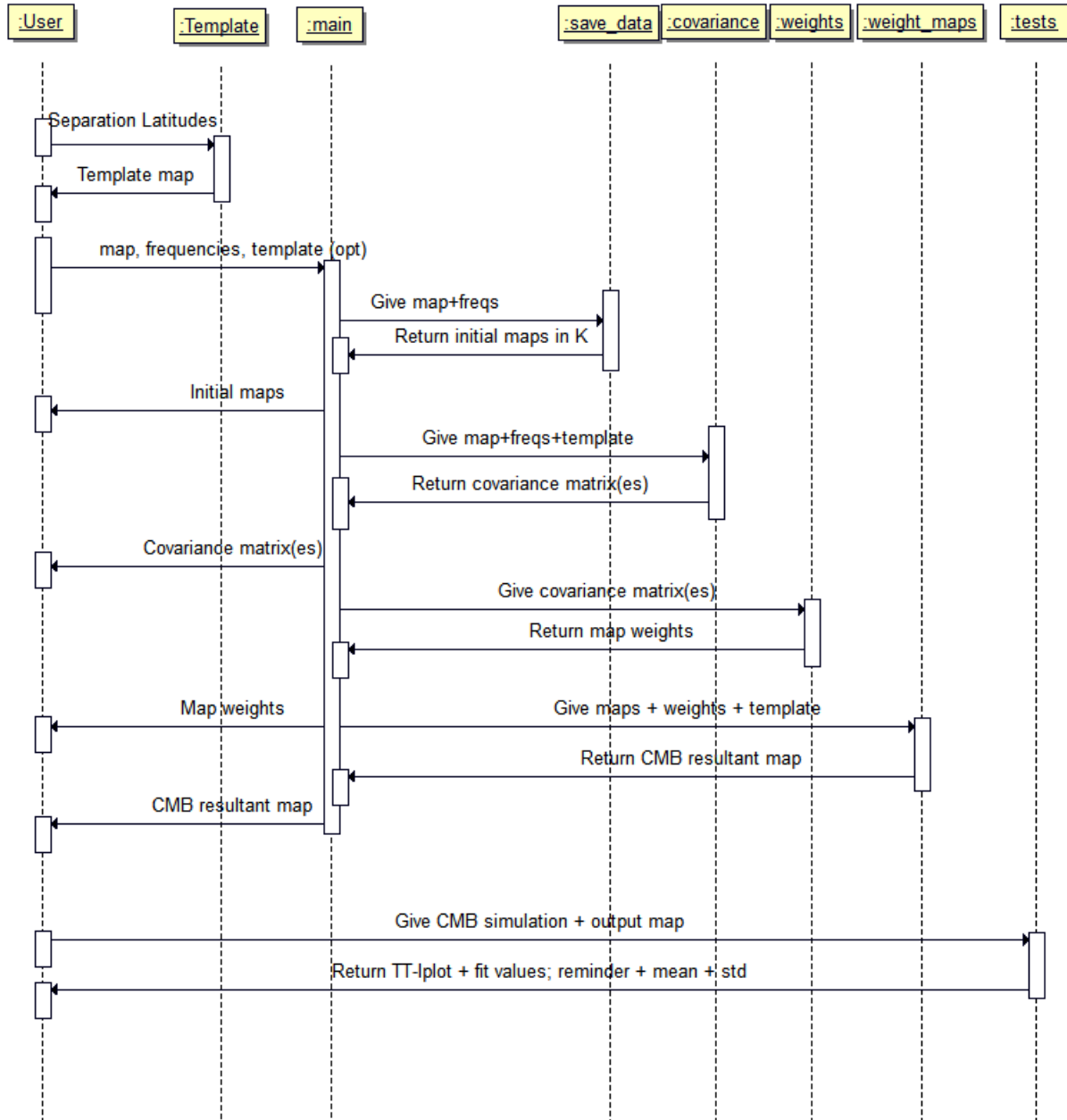


Figura 2.18: *sequence chart* que describe la estructura del programa

Este se divide en tres partes independientes.

La primera, opcional, crea una plantilla del mapa dividido en regiones, que se usará en las siguientes secciones para realizar la separación en las mismas, en lugar de aplicar el algoritmo a todo el mapa de una vez, en caso de que no se use ninguna plantilla de separación. Además, de

cara a la implementación, esta , además de mejorar la eficiencia, sirve para ahorrar trabajo de *debugging*, pudiendo contrastar visualmente dónde se está aplicando esta separación.

La segunda parte corresponde al algoritmo de separación en sí. Tal como se describe esquemáticamente en la Figura 2.18, lo primero es, por comodidad, guardar todos los datos de entrada en una misma carpeta, realizando los cambios de unidades necesarios, de forma que el formato sea el necesario para realizar la separación, así como sumar todos los mapas, obteniendo así el mapa inicial. Una vez con ello, se aplican las ecuaciones descritas en 1.5 (cálculo de la matriz de covarianza y pesos), guardando el resultado en cada paso, tanto para realizar comprobaciones en caso de que sea necesario, así como para ahorrar tener que ejecutar todo el código de nuevo en caso de error. Por último, se aplican los pesos calculados a los mapas, obteniendo el mapa resultante del *CMB*.

La tercera y última parte son las comprobaciones de la validez del resultado, comparándolo con la simulación del *CMB*. Tras una primera comprobación visual (no tiene sentido realizar pruebas sobre un mapa que no tiene absolutamente nada que ver con el resultado esperado), las dos pruebas cuantitativas usadas son un *TT-plot* y el mapa de residuos (otras pruebas que se realizaron en los resultados de Planck implican el análisis del espectro de los mapas. Como en este caso solo se obtiene un resultado suma, no se puede obtener un espectro [10]). Ambos se basan en que, en una situación ideal, los píxeles del resultado deberían tomar el mismo valor que sus píxeles correspondientes en la simulación. El primero es un ajuste de los valores del resultado frente a la simulación que, en teoría, es una recta que pasa por el origen a 45° (es decir, una recta de la forma $x = y$); mientras que el segundo es el mapa de las diferencias (*CMB* resultante - *CMB* simulado), evaluando su media y su desviación estándar.

3 Resultados

Tras definir el formato que va a tener el código, el siguiente paso es su desarrollo en sí. Este se puede descargar en el repositorio de *Github* [40], cuya distribución y forma de instalación se especifica en el apéndice. En esta misma también se encuentran distintas versiones del programa, sin estar completo y con fallos, ya que es útil tener guardado en *Github* la evolución del desarrollo del código, pudiendo acceder a versiones antiguas en caso de que sea necesario (que es una de las funciones principales de git).

Para facilitar el seguimiento de estos resultados, se dividirá la explicación en las secciones que se han desarrollado en 2.5.

3.1 Plantilla de separación

Como se explicó en 1.3, la contaminación no es la misma en todas las regiones (destacando el plano galáctico), por lo que es conveniente aplicar el método por regiones, para lo cual se ha creado una plantilla de separación.

La Figura 3.19 muestra esta plantilla, en la que, tomando como base el formato que siguen los mapas de HEALPix, se ha asignado un valor a cada región un valor, usando el mismo para las partes superior e inferior, que tienen las mismas propiedades.

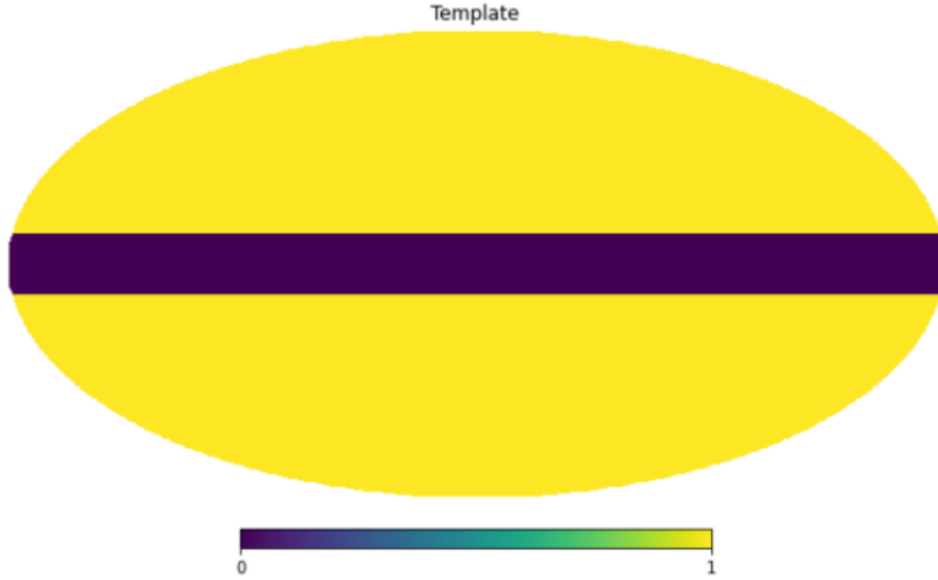


Figura 3.19: Plantilla usada para separar el mapa en regiones, abarcando la región central (en morado) latitudes entre -10° y 10° .

3.2 Sepración de componentes

Antes de poder siquiera representar el mapa inicial, se debe garantizar que la ecuación (18) es correcta, ya que en caso contrario este no tendría sentido. Para ello, se han usado unas simulaciones del *CMB* a distintas frecuencias, proporcionadas por *PLA* [23], que siguen el mismo formato de unidades que las observaciones y, dado que sus valores son bien conocidos, queda bien claro si esta transformación es correcta o no.

El método se ve muy negativamente afectado por fuentes con dependencias espectrales en la posición, que aparecerían en una simulación completa. Es por ello, además de para simplificar el problema, que se ha optado por usar un mapa simplificado, que contiene, además del *CMB*, las fuentes explicadas en 1.3 (sincrotrón, libre-libre y polvo). Estas observaciones han sido construidas sumando los mapas de cada una de las fuentes, proporcionados por *PLA* [23], a cada frecuencia de observación. Como ejemplo, la Figura 3.20 muestra la simulación de observación simplificada a 217 GHz. Además de este, las simulaciones consisten en un total de 9 mapas a distintas frecuencias, entre 30GHz y 857GHz (las frecuencias concretas, así como las características de los instrumentos usados para las mismas, se pueden comprobar en la página del PLA [23]).

Es importante recordar, como se comentaba en 2.4, que se deben igualar las resoluciones de todos los mapas para poder trabajar con todos ellos. La resolución tomada ha sido de 1° .

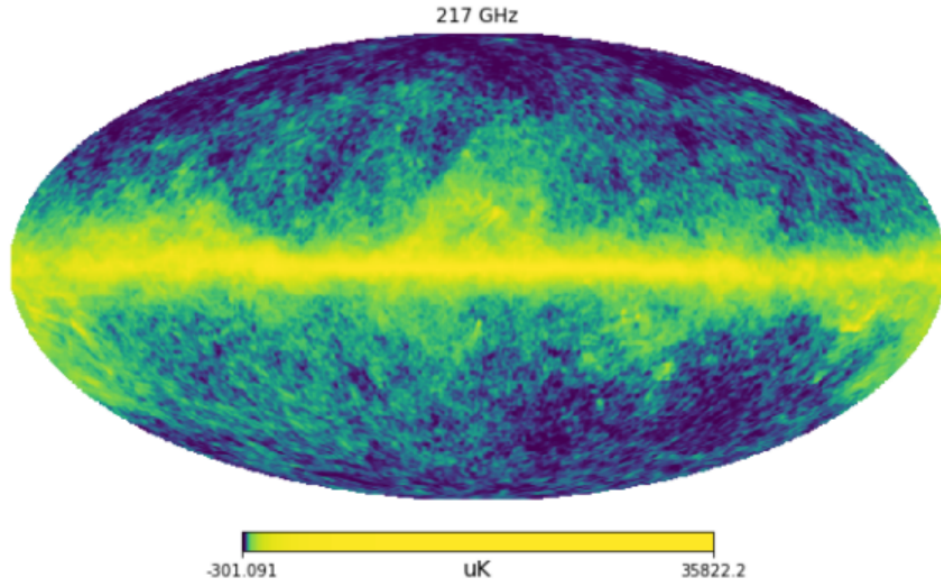


Figura 3.20: Simulación de la observación simplificada (solo con las componentes descritas en 1.3) a 217 GHz.

Tras aplicar la separación a estas simulaciones, se obtienen las Figuras 3.21 y 3.22, haciéndolo a todo el cielo de una vez y dividiéndolo en las regiones mostradas en la Figura 3.19, respectivamente.

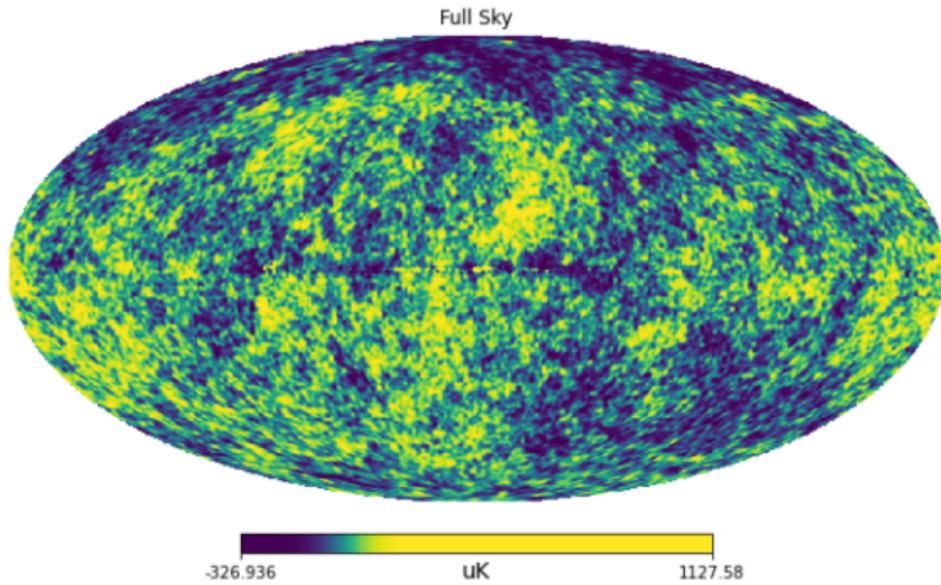


Figura 3.21: Mapa del resultado de aplicar la separación de componentes a todo el cielo de una vez a las observaciones simplificadas, en μK

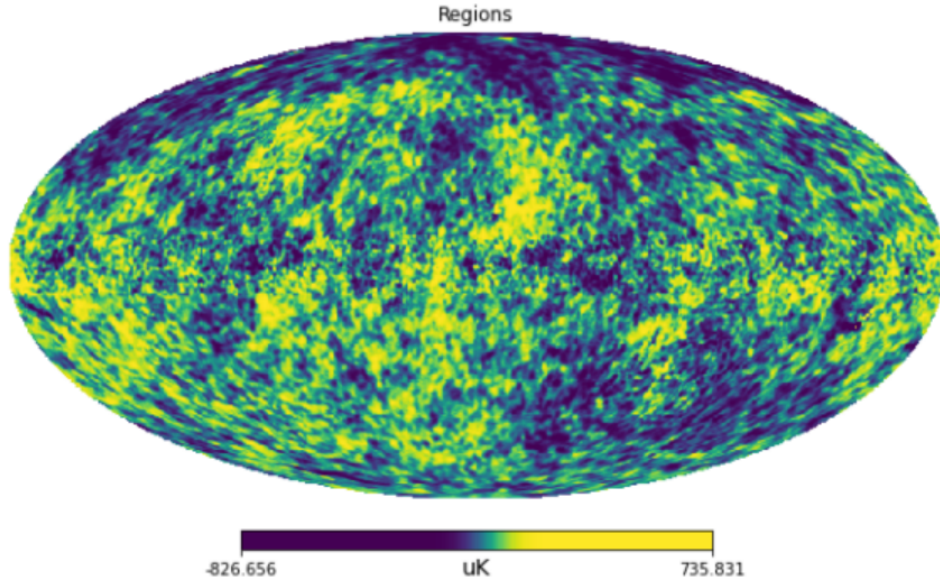


Figura 3.22: Mapa del resultado de aplicar la separación de componentes, dividiendo el cielo según las regiones mostradas en la Figura 3.19, a las observaciones simplificadas, en μK

Tal como se anticipó antes, hay una parte del plano galáctico que no es posible separar del todo, provocando que el abanico de temperaturas sea muy diferente a la esperada (es decir, la temperatura en esa zona proviene, en su mayoría, de los *foregrounds* que no se han quitado del todo, que tienen valores de temperatura órdenes de magnitud más grandes que la del *CMB*. Por lo tanto, el rango de fluctuación mostrado corresponde a esta zona, y no al resto como se desea). Es por ello que, para evitar esta contaminación, generalmente se usan máscaras que tapan estas zonas, pudiendo realizar el trabajo de análisis evitando esta región. En este caso concreto, la máscara usada para realizar estudios de mapas de intensidad del *CMB* se muestra en la Figura 3.23.

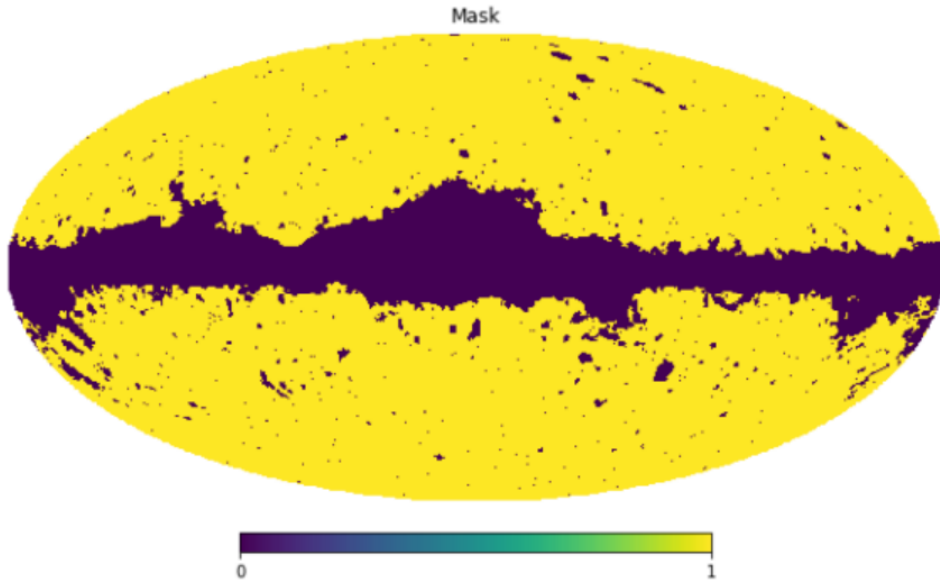


Figura 3.23: Máscara usada para el análisis de mapas de intensidad del CMB. Las regiones son las que se tapan, conservando las amarillas. [23]

Cuando se usa esta máscara, los mapas resultantes son las Figuras 3.24 y 3.25. Se observa como, en este caso, el rango de temperaturas de los mapas entra dentro de los valores esperados.

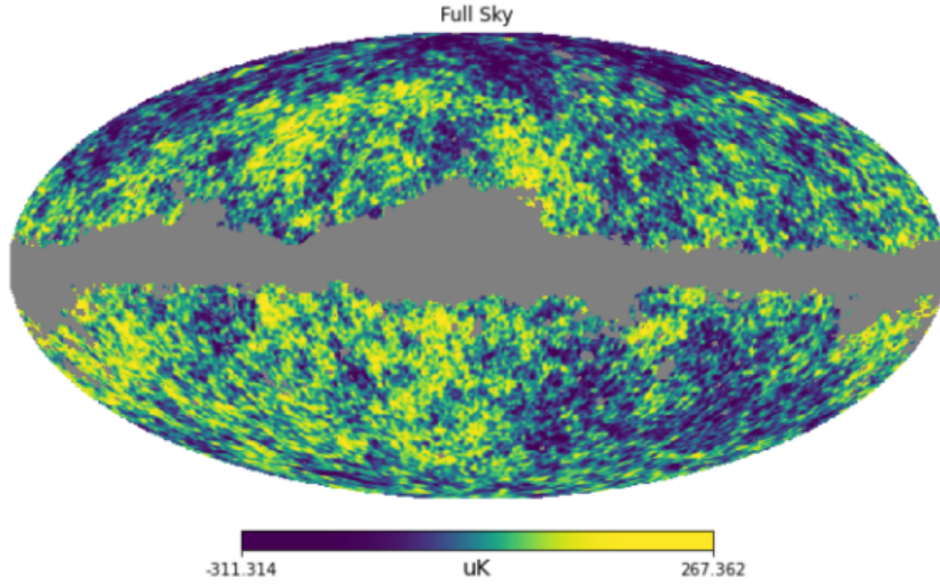


Figura 3.24: Mapa del resultado de aplicar la separación de componentes a todo el cielo de una vez a las observaciones, usando la máscara de intensidad, en μK

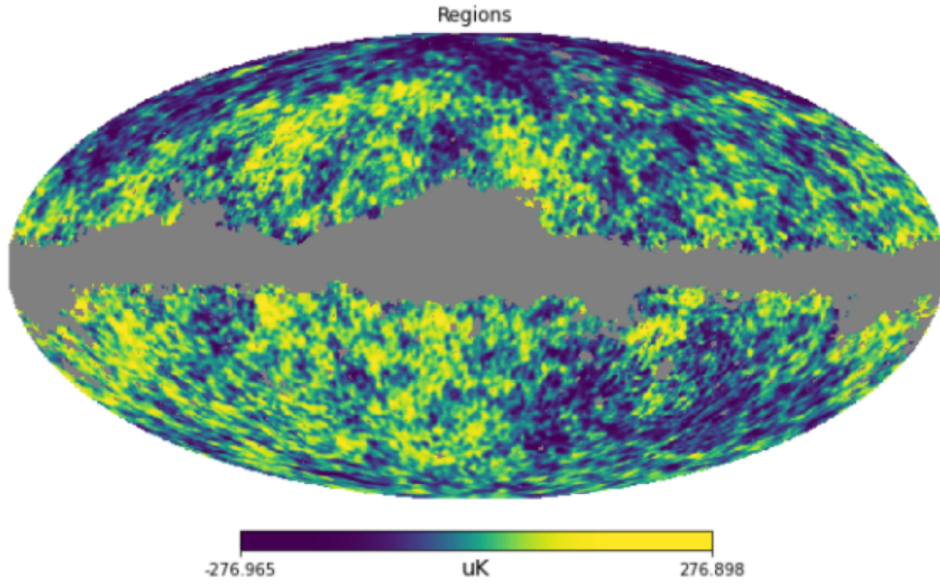


Figura 3.25: Mapa del resultado de aplicar la separación de componentes, dividiendo el cielo según las regiones mostradas en la Figura 3.19, a las observaciones simplificadas, usando la máscara de intensidad, en μK

3.3 Comparación de los resultados con el *CMB*

Para analizar los datos resultantes, lo primero es tener un mapa del *CMB* aislado con el que compararlo, proporcionado por *PLA* [23]. Este se muestra en la Figura 3.26.

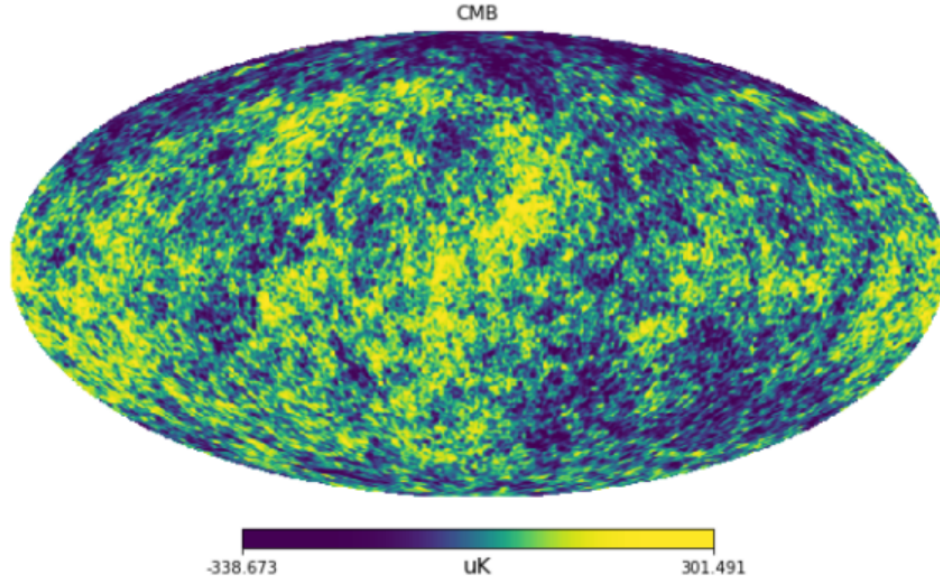


Figura 3.26: Mapa del *CMB* aislado. [23]

Como se comentó previamente, las dos pruebas que se han realizado sobre los resultados (además de estudiar sus parámetros estadísticos) son el diagrama TT-plot y el mapa de residuo.

Para comenzar, los valores de media y desviación estándar de los tres mapas individualmente se recogen en la Tabla 3.3.

Lo más llamativo de la misma es que, observando los valores de desviaciones estándar σ , resulta ser más baja en los mapas resultantes que en la propia simulación. Esto no tiene sentido, ya que tomando como punto de partida la ecuación (9), la desviación estándar del resultado es la suma de las del mapa del *CMB* (Figura 3.26) y las de los *foregrounds* restantes (en este caso no se ha usado ruido). Dado que todas esas cantidades son positivas, en el mejor de los casos serán iguales a las del mapa original, pero en ningún caso menores. De todas formas, al no ser la diferencia muy grande puede que sea justificado con los errores de las próximas pruebas

Tabla 3.3: Valores de la media y desviación estándar, tanto del mapa del CMB, como de los resultados aplicando el método a todo el cielo completo (*full sky*) y dividido por regiones (*regions*). La parte superior son los valores sin usar la máscara, mientras que en la inferior se ha empleado la misma

	Media(μ K)	σ (μ K)
cmb	0.171	76.837
full sky	-8.905	69.876
regions	-6.062	66.196
cmb	-8.288	75.872
full sky	-13.5588	68.825
regions	-9.171	64.572

3.3.1 Mapa del residuo

La otra prueba a realizar es el mapa del residuo (diferencia de los mapas resultantes con el *CMB*). Esto lleva a las Figuras 3.27 a 3.30.

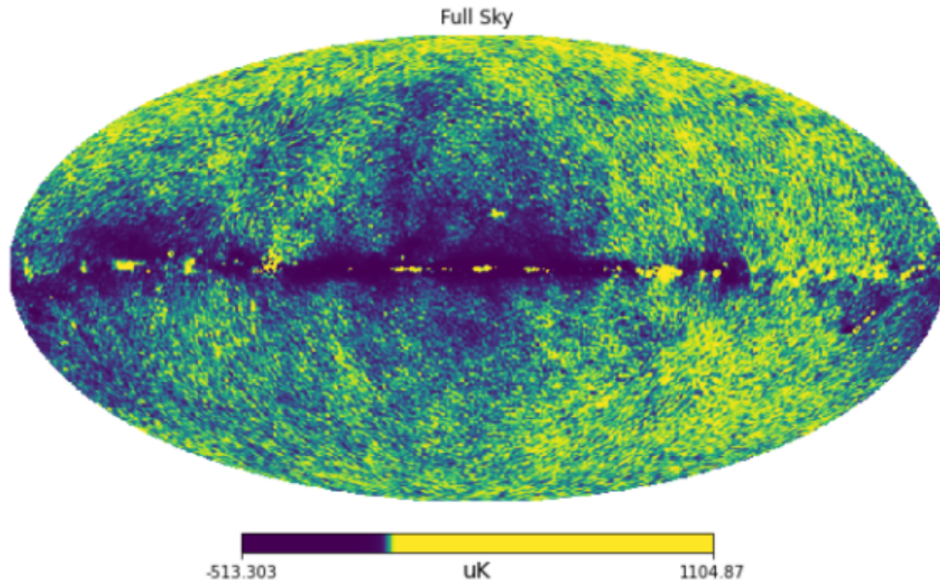


Figura 3.27: Mapa del residuo del resultado obtenido aplicando el método a todo el cielo de una vez, en μK . Su media es de $-9,022\mu\text{K}$, y su desviación estándar, $17,096\mu\text{K}$.

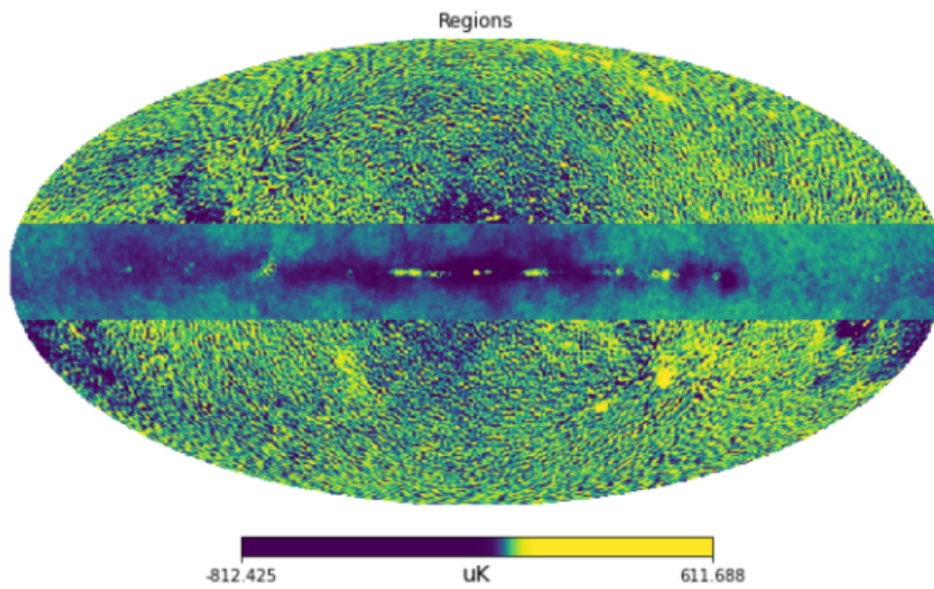


Figura 3.28: Mapa del residuo del resultado obtenido aplicando el método al cielo dividido en regiones, en μK . Su media es de $-6,179\mu\text{K}$, y su desviación estándar, $30,000\mu\text{K}$.

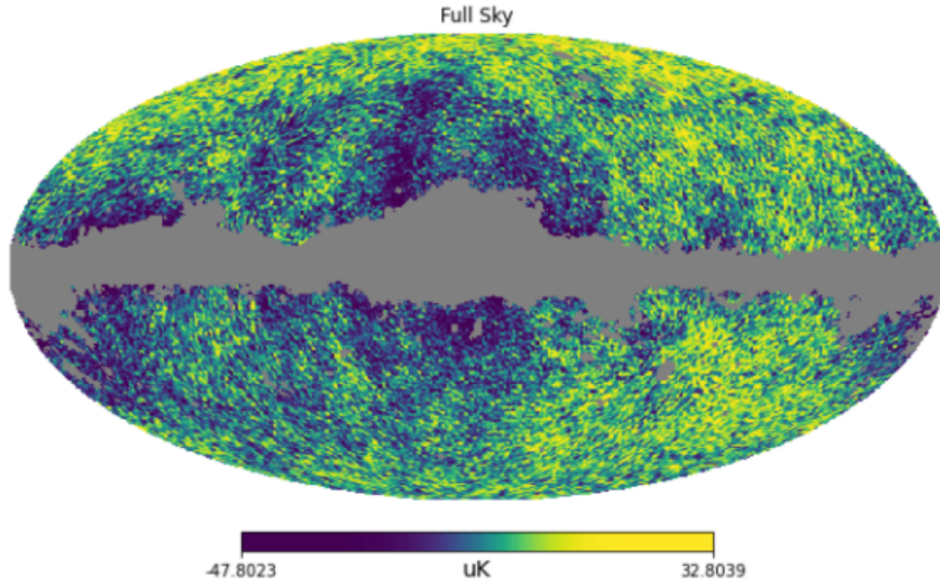


Figura 3.29: Mapa del residuo del resultado obtenido aplicando el método a todo el cielo de una vez, usando la máscara, en μK . Su media es de $-5,270\mu\text{K}$, y su desviación estándar, $8,471\mu\text{K}$.

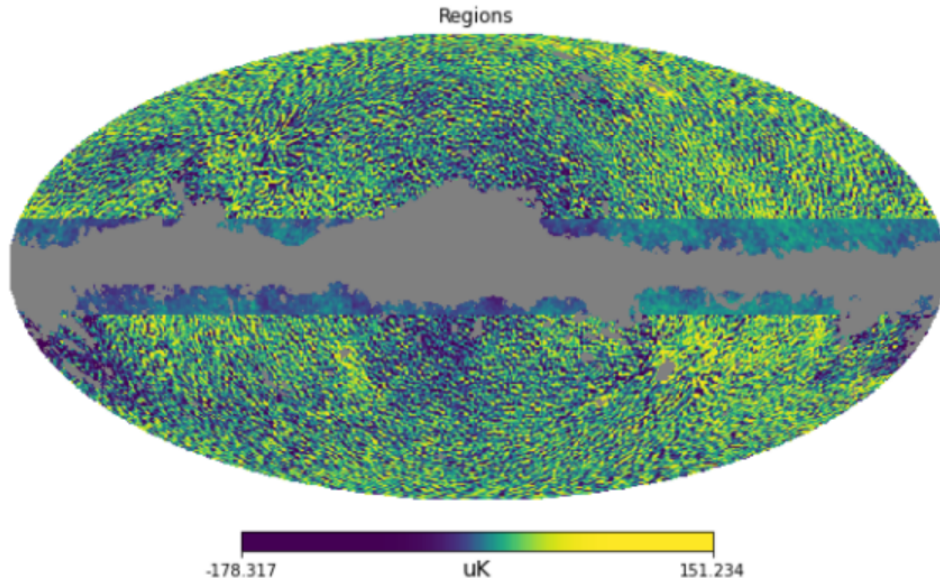


Figura 3.30: Mapa del residuo del resultado obtenido aplicando el método al cielo dividido en regiones, usando la máscara, en μK . Su media es de $-0,873\mu\text{K}$, y su desviación estándar, $27,359\mu\text{K}$.

Antes de hacer un análisis más en detalle de los resultados numéricos, cabe destacar algunos aspectos importantes sobre las imágenes de los mapas. En primer lugar, que es lo que más llama la atención, es la discontinuidad en la Figura 3.28. Como se anticipó, el hecho de separar el cielo por regiones genera una discontinuidad en los límites de las mismas, y se plasma claramente en estos mapas. Por otro lado, destacar que, al comparar las Figuras 3.27 y 3.28, hay una aparente mejoría en la eliminación de la contaminación a medias latitudes, que podría llegar a pensar que la separación mejora en resultado en este caso. Sin embargo, esto puede deberse a un efecto de la escala de colores, que es distinta en los dos mapas. Por último, siguiendo en esta línea de comparación, se observa que, al tratar todo el cielo de una vez, no se elimina de todo bien la contaminación a latitudes medias, quedando un remanente, mientras que al separar el plano galáctico del resto, esto no ocurre, resultando un mapa más limpio en esta región.

Pasando ya a análisis cuantitativos, aunque a primera vista pueda parecer que el resultado de la Figura 3.30 es mejor que el de la Figura 3.29 porque su media es más baja, realmente es al contrario, ya que el error es tres veces mayor. Comparando estos valores con los datos de la Tabla 3.3, se tiene que el error de los mapas es bastante menor que la desviación estándar que el mapa del *CMB* de la Figura 3.26, con un error del 11 %.

3.3.2 *TT*-plot

Si se representa en una gráfica los resultados de las Figuras 3.21 y 3.22, se obtienen las Figuras 3.31 y 3.32.

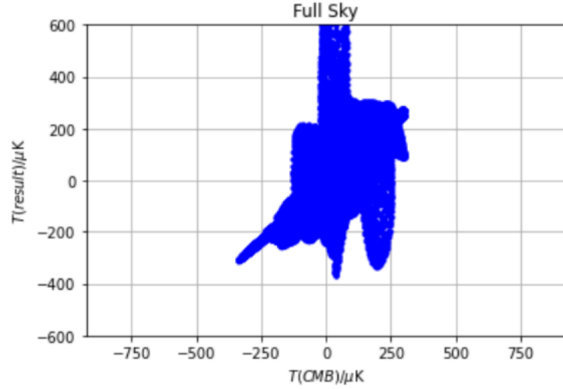


Figura 3.31: Temperatura del mapa resultante al aplicar el método a todo el cielo de una vez, frente a la temperatura del mapa del *CMB* aislado (Figura 3.26), sin usar la máscara.

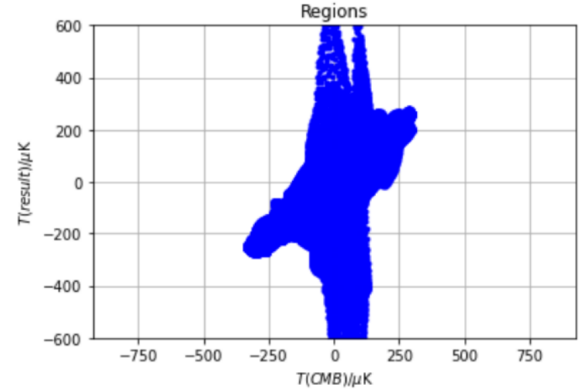


Figura 3.32: Temperatura del mapa resultante al aplicar el método al mapa dividido en regiones, frente a la temperatura del mapa del *CMB* aislado (Figura 3.26), sin usar la máscara.

Al igual que pasaba con las imágenes de los mapas y los residuos, la contaminación del plano galáctico hace que sea necesario el uso de la máscara. Con ello, y haciendo un ajuste lineal, se obtienen las Figuras 3.33 y 3.34.

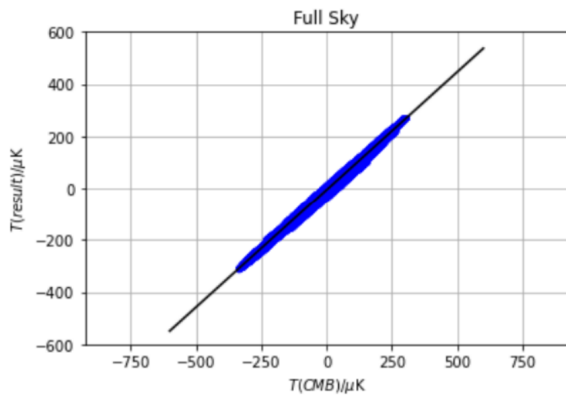


Figura 3.33: Temperatura del mapa resultante al aplicar el método a todo el cielo de una vez, frente a la temperatura del mapa del *CMB* aislado (Figura 3.26), usando la máscara. El resultado del ajuste es $y = (0,905205 \pm 0,000019)x - (6,0553 \pm 0,0014)\mu\text{K}$.

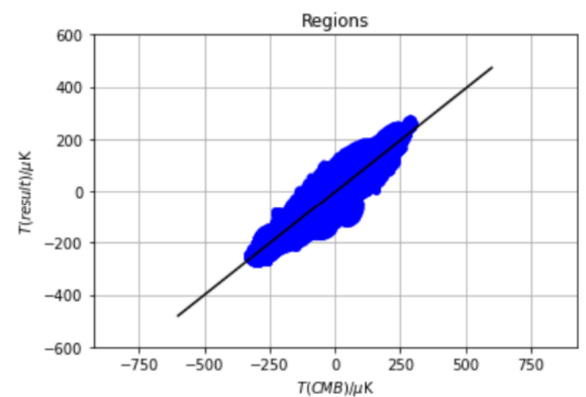


Figura 3.34: Temperatura del mapa resultante al aplicar el método al mapa dividido en regiones, frente a la temperatura del mapa del *CMB* aislado (Figura 3.26), usando la máscara. El resultado del ajuste es $y = (0,79472 \pm 0,00009)x - (3,841 \pm 0,007)\mu\text{K}$.

Aquí se observa claramente el por qué, al aplicar el método por regiones, aumenta el error. Comparando las Figuras 3.33 y 3.34, mientras que la primera se ajusta casi perfectamente a una recta, la segunda tiene una dispersión mayor. Además, la pendiente, que en teoría tiene que ser 1, tiene un valor más alejado del esperado en este caso. Por último, la ordenada en el origen de la segunda gráfica algo más próxima a cero que la primera. Esto significa que, en promedio, a pesar de la dispersión, los resultados de la Figura 3.34 se aproximan más a los del *CMB* que los de la Figura 3.33. Sin embargo, dado que el error es de $8,471\mu\text{K}$ (obtenido a partir de los residuos), esta diferencia entra dentro de los márgenes de error. Es decir, separar el cielo por regiones empeora el resultado (al menos, en este caso).

3.4 Análisis de los resultados

Una vez con todos los resultados calculados, el siguiente paso es realizar su análisis.

Comenzando por los parámetros de la Tabla 3.3, se observa que los valores resultantes son muy similares a lo esperado en ambos casos. En teoría, el valor medio del mapa, dado que son fluctuaciones en torno a la temperatura media del *CMB*, debería ser de 0K , obteniendo resultados del orden del μK , con una dispersión de 10^{-5}K , del mismo orden de magnitud que el mapa con el que se está comparando. Como se comentó, en 3.3, lo más extraño es que salgan valores más bajos que los originales, pero esto se puede justificar con los errores de los otros análisis. Tomando como referencia el mejor de los resultados, la pendiente de la Figura 3.33 es de $0,905205 \pm 0,000019$, $9,48\%$ menor del valor esperado. Tomando este error como referencia, esta bajada de la desviación estándar entra dentro de lo esperado (haciendo números exactos, el límite inferior con este margen de error es de $69,553\mu\text{K}$, por debajo del valor resultante). En el caso del resultado por regiones, el error crece a un $20,53\%$, resultando el límite inferior $61,064\mu\text{K}$, también por debajo del resultado. En ambos casos esta discrepancia es coherente con su error.

Por otra parte, en los mapas de residuos se obtienen valores medios próximos a cero (como se esperaba), y con un error un 11% más bajo que la desviación estándar del *CMB*. Por un lado esto es positivo, porque quiere decir que se ha llegado a un mapa con un error del 11% , pero para poder hacer cosmología se necesita bajar este hasta, al menos, un 1% . Es decir, el método funciona y consigue resultados razonables, pero no es lo suficientemente potente para trabajar con los datos de *Planck*.

En cuanto a los *TT*-plot, los ajustes distan bastante de lo esperado, que es una recta a 45° , o una pendiente de valor 1 (el píxel en el mapa del *CMB* tomando un valor similar al correspondiente en el mapa resultante), mientras que las pendientes obtenidas han sido $0,905205 \pm 0,000019$ y $0,79472 \pm 0,00009$, en la separación a cielo completo y por regiones, respectivamente. Esto quiere decir que el resultado de la Figura 3.33 se aproxima bastante mejor a los valores esperados que el de la Figura 3.34. En cuanto a las ordenadas, como ya se comentó en 3.3.2, aunque parezca que en el segundo caso de media se aproxime más a los valores esperados, tomando los márgenes de error de los residuos sigue siendo mejor el resultado en el primer caso.

3.5 Aplicación al mapa completo

Una vez hecho el estudio del mapa parcial, es conveniente contrastar qué ocurre al aplicar este mismo método a un mapa que contiene más *foregrounds*, aproximándose así más a la realidad. Dichos *foregrounds* son [25]:

-**Efecto Sunyaev-Zel'dovich térmico** (*thermalsz*): variaciones de la energía de los fotones emitidos por el *CMB* por colisiones con electrones muy energéticos del medio intracumular.

-**Fuentes puntuales extragalácticas** (*strongirps*, *faintirps*, *strongradiops*, *faintradiops*): son cuatro tipos de fuentes puntuales extragalácticas (de ahí la terminación *ps*, viene de 'point source'), que se diferencian en cómo se han simulado. Dos de ellas son en infrarrojo (*ir*, *infrared*) y las otras en radio. El primer identificador (*strong* y *faint*) determina la intensidad de la fuente. Las primeras, si la fuente es muy brillante se colocan en su posición real, introduciendo sus características (brillo, color, ...). Por otro lado, las segundas, más débiles, se colocan al azar, siguiendo distribuciones teóricas de modelos de población.

-**Fondo infrarrojo lejano** (*Far Infrared Background*, o *firb*): Una gran cantidad de fuentes emiten en el infrarrojo lejano, siendo demasiado tenues para verse individualmente. La suma de todas ellas forma, en conjunto, este *foreground*.

El resultado en este caso es la Figura 3.35

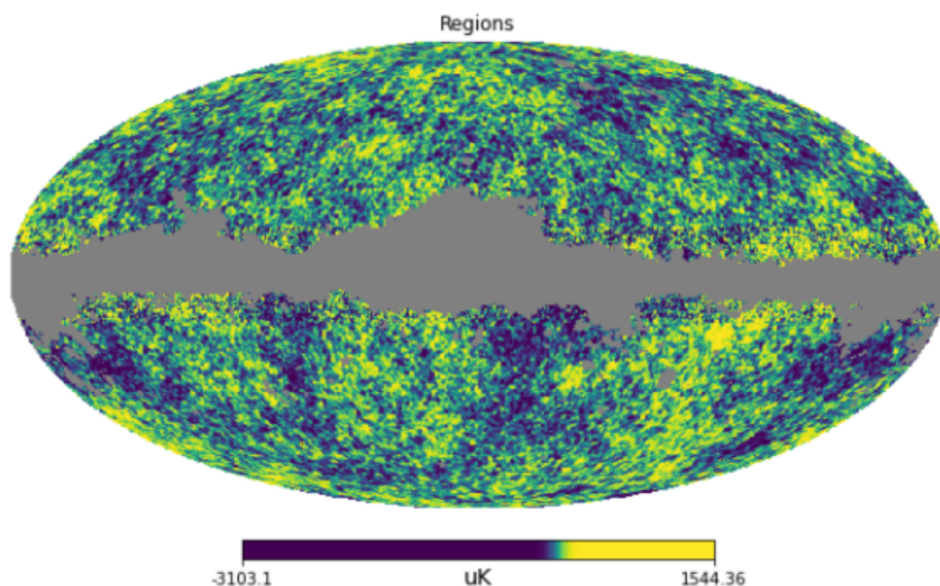


Figura 3.35: Resultado de aplicar el método a la simulación que incluye todos los *foregrounds*, dividiendo el mapa por regiones, usando la máscara, en μK

Se observa que el intervalo de temperaturas difiere mucho del previamente obtenido. Por lo tanto, el método no sirve en todos los casos, habiendo algún componente que no es capaz de separar.

4 Conclusiones

4.1 Valoración de los resultados

En primer lugar, destacar que el objetivo del trabajo, tal como se comentó en 1.6, no era obtener un resultado tan preciso como el de *Planck*, sino comprobar hasta que punto el *ILC* da resultados satisfactorios, al tiempo que se aprende a trabajar con el formato de datos que se usa en el campo de la astrofísica.

Teniendo en cuenta esto, los resultados son relativamente positivos. Se ha visto que es mucho más eficaz al aplicarlo al mapa completo, siempre que se use la máscara, tanto por la discontinuidad como la dispersión que se generan al aplicarlo por regiones. Una de las grandes ventajas de

usar este método es que el proceso completo de separación es muy rápido teniendo en cuenta las dimensiones de los datos empleados (cada mapa contiene más de 12 millones de píxeles). Tras un proceso de optimización, esencial en estos casos, se ha conseguido una versión del programa que crea el mapa resultante en apenas un minuto, siendo la parte más lenta la parte de los ajustes, en la cual no se puede hacer demasiado.

Por otra parte, comparando el resultado con el obtenido por la propia misión, como puede ser el mapa de SEVEM de la Figura 1.16, se observa que ambos mapas son muy similares. La mayor diferencia entre ambos posiblemente sea el remanente del plano galáctico que, aunque también está presente en el resultado oficial, es mucho menos intenso que en el obtenido. Aunque al usar la máscara esto no sea un problema muy grande de cara al análisis, es conveniente tratar de minimizar la contaminación en esta región, que pudiera llegar a influir negativamente en otras zonas (ver Figura 3.33).

Siguiendo en la línea del problema del plano galáctico, se ha comprobado que el hecho de separar el mapa en regiones a las que aplicar el método de forma independiente contribuye positivamente al resultado, pero los errores aumentan mucho. Como se dispone de la máscara de intensidad, lo que ocurra en esta región no tiene mucha relevancia en el resultado (ni los códigos oficiales de separación de *Planck* son capaces de realizar esta separación).

Por último, se ha comprobado que este método, al asumir muchas cosas acerca de los *foregrounds*, no puede separar todos ellos, es decir, no serviría en la realidad. En concreto, las fuentes que dependen de la posición hacen que el método falle. En este caso, las fuentes puntuales mencionadas en 3.5, como su propio nombre indica, cumplen este requisito.

4.2 Opciones de mejora

El método empleado es muy sencillo, por lo que el margen de mejora de los resultados es enorme. Por supuesto, la primera prueba a hacer es ver qué ocurre al incluir ruido (no se ha añadido para hacer una primera prueba sencilla, así como las limitaciones de espacio de almacenamiento). También se podría ampliar la estadística, usando un gran número de simulaciones sobre las que aplicar el código, de forma que se aumentaría la precisión de los resultados. Esto no se ha hecho por falta de tiempo y espacio de almacenamiento, ya que estas simulaciones pueden llegar a pesar varios gigas, de los que no se dispone, y la función de suavizado es muy lenta, no disponiendo del tiempo necesario para obtener varios mapas.

Por otro lado, el propio método es mejorable si se usa la matriz de mezcla, A , que contiene, en esencia, la información de la Figura 1.10 (es una matriz $n \times n$, siendo n el número de componentes totales del mapa, cuyos elementos son los productos de las funciones de intensidad de cada uno de ellos). Con esto, partiendo de que cada píxel es una combinación de lineal de n fuentes independientes, se llega a

$$D = AS + N \quad (19)$$

Donde D es la matriz de datos completa, S la matriz de componentes, y N el ruido.

La gran ventaja de esto es que, además de mejorar la precisión de los resultados, ya que se está introduciendo explícitamente información característica cada uno de los componentes presentes en la imagen, no solo se recupera una imagen del *CMB*, sino que también se obtiene un mapa de cada *foreground*.

Los principales problemas, además de la propia dificultad de programación, son que el ruido tiene que ser previamente eliminado para que los datos contengan únicamente la información de los componentes incluidos en la matriz de mezcla (lo cual no es muy complejo, pero hay que hacerlo); se necesita conocer qué componentes están presentes en la imagen *a priori*, no pudiendo ni siquiera comprobar el método con el mapa completo (además de que a mayor número de componentes, más mapas se requieren, por lo que no es muy eficaz aplicarlo con un gran número de fuentes presentes en el mapa), y, además, se debe asegurar de que las funciones que describen su perfil de intensidad, además de hacerlo correctamente, estén bien calibradas para que su producto pueda describir su mezcla. Es por todas estas dificultades que se ha optado por usar la versión simplificada, obteniendo igualmente resultados satisfactorios.

4.3 Resultados del aprendizaje

Durante la carrera he realizado algunos experimentos dentro del campo de la astrofísica, pero por lo general siempre se trabajaban con pequeñas cantidades de datos y con un método sencillo que venía dado de antemano (a excepción de un par de trabajos algo más grandes, cuya dificultad de programación era similar a la de este proyecto, pero igualmente se trabajaba con pocos datos). A diferencia de los anteriores, en su mayoría experimentos realizados a mediados del siglo pasado, la misión *Planck* se terminó hace apenas cuatro años a fecha de hoy, de manera que tengo una mejor idea del contexto de trabajo actual en el campo de la astrofísica (al menos, una parte). Con respecto al método matemático, no se trataban de dos o tres ecuaciones simples, proporcionadas por el profesor, sino que he tenido que leer varios artículos sobre el método, contrastando distintas formas en las que se expresan las ecuaciones, seleccionando las adecuadas para este caso concreto, teniendo que desarrollarlas en algunos casos, tras lo que tenía que realizar pruebas con ellas, modificándolas hasta encontrar la versión que necesitaba. Todo ello ha supuesto experiencia en resolución de problemas, cuya solución no es conocida de antemano y se deben plantear varias formas de llegar a una de las posibles soluciones válidas, habilidad de vital importancia en el mercado laboral actual. Durante todo el trabajo no había hecho el suavizado que se mencionaba en 3.2 por un fallo de comunicación, de forma que los resultados no salían bien. Esto ha reforzado mi pensamiento crítico, haciendo que tenga que buscar, a través de todas las pruebas que fueran surgiendo, donde estaba el error, más productivo que el acercamiento en experimentos pequeños en los que concluyes que está mal y no se mira más allá. Aunque finalmente no haya llevado a ningún sitio, la experiencia ha resultado extremadamente positiva.

Con respecto a la programación, hasta ahora todos los códigos que había desarrollado solo los revisaba yo y, como mucho, una persona más. Dado que, en este caso, más gente va a tener acceso al mismo, he cuidado los comentarios durante el desarrollo, tanto para que sepan tanto utilizar el código como saber corregir posibles *bugs*, en caso de que sea necesario. Además de que, siendo un trabajo de varios meses de duración, incluso al programador le vienen los mismos para entender el funcionamiento de partes del código en las que no ha trabajado durante un largo periodo de tiempo. Además, por estos mismos motivos, tener un código limpio y bien organizado es también esencial. Gracias a esto he mejorado algo mis habilidades en UML, así como la herramienta BOUML, un lenguaje que sirve para diseñar programas antes de comenzar a desarrollarlos, que permite tener claro de antemano la estructura que va a tener el código (especialmente útil para grandes programas que implican a varias personas trabajando, que necesitan que cada sección encaje con las otras realizadas por sus compañeros), lo que acelera el desarrollo, ayuda a tener un código más limpio, que facilita la búsqueda y corrección de fallos en el mismo (*debugging*).

A esto se le suma, como se ha explicado anteriormente, lo crucial que es la optimización del código. Como anteriormente he trabajado con pocos datos (siendo 100 valores una gran cantidad, para tener una idea de la escala), un código subóptimo no supone un problema, ya que, en el

peor de los casos, se pierden unos pocos segundos. Cuando se usa un gran volumen de datos este problema se acentúa, de forma que he tenido que realizar un trabajo de optimización que ahorra varios minutos en cada ejecución (principalmente eliminando bucles, que consumen mucho tiempo, por funciones alternativas más rápidas), lo que permite hacer más pruebas.

Empezando a particularizar en este trabajo concreto, siguiendo en la línea del volumen de los datos, cuando se trabaja con pocas decenas de ellos la forma de almacenamiento es indiferente, pudiendo incluso crear los ficheros de datos a mano en algunos casos. Sin embargo, con una cantidad grande esto no es viable, tanto en espacio de almacenamiento como en tiempo. El formato FITS, que encripta los datos, es muy eficiente en ambos aspectos. Durante este trabajo he aprendido a manejar ficheros .fits a través de la librería *astropy*, que permite, además de leer los datos, acceder a los encabezados que describen los mismos, sin los cuales estos datos son simplemente números sin ningún significado físico.

En lo que se refiere a la física que hay detrás del trabajo, he comprendido mejor la importancia que tiene el *CMB* en los campos de la astrofísica y la cosmología, al tiempo que visualizo y entiendo cuales son los distintos procesos físicos responsables de la emisión de radiación en el universo (al menos, algunos de los más importantes), y la importancia que tiene conocer bien cada uno de ellos para poder llevar a cabo un algoritmo de separación de componentes y poder estudiar otras fuentes de forma independiente (de la misma forma que se ha hecho en este trabajo con el *CMB*).

Referencias

- [1] Lee Mohon. *Sagittarius A*: NASA Telescopes Support Event Horizon Telescope in Studying Milky Way's Black Hole*. https://www.nasa.gov/mission_pages/chandra/images/sagittarius-a-nasa-telescopes-support-event-horizon-telescope-in-studying-milky-ways.html. Accedido 07/09/2022. Mayo de 2022.
- [2] *EHT Collaboration*. https://en.wikipedia.org/wiki/File:EHT_Sagittarius_A_black_hole.tif. Accedido 07/09/2022.
- [3] Getsemary Báez. *History of Radio Astronomy*. https://swaves.gsfc.nasa.gov/pdf_files/History%20of%20Radio%20Astronomy.pdf. Accedido 07/09/2022.
- [4] Brooke Boen. *Supermassive Black Hole Sagittarius A**. https://www.nasa.gov/mission_pages/chandra/multimedia/black-hole-SagittariusA.html. Accedido 07/09/2022. Ago. de 2017.
- [5] *The History of Radio Astronomy*. <https://public.nrao.edu/radio-astronomy/the-history-of-radio-astronomy/>. Accedido 07/09/2022.
- [6] *NRAO/AUI/NSF*. <https://public.nrao.edu/gallery/karl-jansky-and-his-merrygoround/>.
- [7] https://es.m.wikipedia.org/wiki/Archivo:Green_Banks_-_Grote_Reber_Radio_Telescope.jpg.
- [8] *The Most Important Equation in Astronomy!* <https://spacemath.gsfc.nasa.gov/moon/6Page50.pdf>. Accedido 10/08/2022.
- [9] *Astronomical interferometer*. https://en.wikipedia.org/wiki/Astronomical_interferometer. Accedido 07/09/2022.
- [10] S. M. Leach y col. "Component separation methods for the Planck mission". En: *Astronomy Astrophysics* (2008). DOI: 10.1051/0004-6361:200810116.
- [11] NASA. *Ask an Astrophysicist*. Accedido 28/08/2022.
- [12] ESA. https://www.esa.int/var/esa/storage/images/esa_multimedia/images/2013/03/planck_history_of_universe/12584080-5-eng-GB/Planck_history_of_Universe.jpg.
- [13] *Radiation Blackbody*. <http://rossby.msrc.sunysb.edu/~marat/MAR542/ATM542-Chapter2.pdf>. Accedido 29/08/2022.
- [14] *Test of Big Bang: The CMB*. https://wmap.gsfc.nasa.gov/universe/bb_tests_cmb.html. Accedido 07/09/2022.
- [15] *Black-body radiation*. https://en.wikipedia.org/wiki/Black-body_radiation. Accedido 07/09/2022.
- [16] Gourav Khullar. 'One Direction' - Isotropic Universe or not? <https://astrobites.org/2016/09/29/one-direction-isotropic-universe-or-not/>. Accedido 07/09/2022. Sep. de 2016.
- [17] NASA. *What is the Inflation Theory?* https://wmap.gsfc.nasa.gov/universe/bb_cosmo_infl.html#:~:text=The%20Horizon%20Problem%3A,the%20age%20of%20the%20universe.. Accedido 27/08/2022.
- [18] APPEL News Staff. *This Month in NASA History: COBE Scientists Were Celebrated for the CMB*. <https://appel.nasa.gov/2016/10/20/this-month-in-nasa-history-cobe-scientists-were-celebrated-for-the-cmb/>. Accedido 10/08/2022. Oct. de 2016.
- [19] *metsahovi*. http://www.metsahovi.fi/quasar/planck/images/S_foreground.gif. Accedido 08/09/2022.

- [20] ESA/Hubble. <https://esahubble.org/images/heic1208a/>. Accedido 27/08/2022.
- [21] Jacques Delabrouille y J. Cardoso. “Diffuse Source Separation in CMB Observations”. En: *Lecture Notes in Physics* 665 (mar. de 2007). DOI: 10.1007/978-3-540-44767-2_6.
- [22] Dorothea Samtleben, Suzanne Staggs y Bruce Winstein. “The Cosmic Microwave Background for Pedestrians: A Review for Particle and Nuclear Physicists”. En: (2008).
- [23] *Planck Legacy Archive*. <https://pla.esac.esa.int/#home>. Accedido 21/08/2022.
- [24] *A Single-Layer Atmosphere Model*. <https://www.acs.org/content/acs/en/climatescience/atmosphericwarming/singlelayermodel.html>. Accedido 06/09/2022.
- [25] *2018 Planck Explanatory Supplement*. https://wiki.cosmos.esa.int/planck-legacy-archive/index.php/Main_Page. Accedido 21/08/2022.
- [26] Alves, Joao y col. “Planck 2018 results”. En: *A&A* 641 (2020), E1. DOI: 10.1051/0004-6361/202039265. URL: <https://doi.org/10.1051/0004-6361/202039265>.
- [27] *Planck*. https://www.esa.int/Enabling_Support/Operations/Planck. Accedido 16/08/2022.
- [28] Marta Seror García. *The legacy of Planck*. <https://www.somma.es/articles/legacy-planck>. Accedido 16/08/2022. Mar. de 2019.
- [29] *ESA*. https://www.esa.int/var/esa/storage/images/esa_multimedia/images/2006/11/planck/9849681-3-eng-GB/Planck_pillars.jpg.
- [30] *The legacy of Planck*. <https://ifca.unican.es/en-us/news/Paginas/The-legacy-of-Planck.aspx>. Accedido 28/08/2022.
- [31] PLA. *2018 CMB maps*. https://wiki.cosmos.esa.int/planck-legacy-archive/index.php/CMB_maps. Accedido 28/08/2022.
- [32] H. Eriksen y col. “On Foreground Removal from the Wilkinson Microwave Anisotropy Probe Data by an Internal Linear Combination Method: Limitations and Implications”. En: *The Astrophysical Journal* 612 (mar. de 2004). DOI: 10.1086/422807.
- [33] C. L. Bennett y col. “First Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Foreground Emission”. En: (jun. de 2003).
- [34] *Anaconda*. <https://www.anaconda.com/>. Accedido 30/08/2022.
- [35] *FITS Documentation*. https://fits.gsfc.nasa.gov/fits_documentation.html. Accedido 21/08/2022.
- [36] Krzysztof M. Górski y col. “The HEALPix Primer”. En: (jun. de 2018).
- [37] NASA. *FITS Keyword Dictionaries*. https://heasarc.gsfc.nasa.gov/docs/fcg/standard_dict.html. Accedido 21/08/2022.
- [38] *HEALPix Formats*. <https://gamma-astro-data-formats.readthedocs.io/en/latest/skymaps/healpix/index.html>. Accedido 21/08/2022.
- [39] NASA. *Fluctuations in the Cosmic Microwave Background*. https://wmap.gsfc.nasa.gov/universe/bb_cosmo_fluct.html#:~:text=The%20actual%20temperature%20of%20the%20cosmic%20microwave%20background%20is%202.725%20Kelvin.. Accedido 29/08/2022.
- [40] Sergio Blanco Gómez. *Planck_ILC*. https://github.com/s3rbg/TFG_ILC. 2022.

Instrucciones de uso del código

Todo el código que forma parte del núcleo del trabajo se puede encontrar en *GitHub*. En primer lugar, mencionar que no es compatible con *Windows* porque *Healpy* no tiene esta compatibilidad. Para descargarlo, simplemente se debe entrar al repositorio de *GitHub* [40] y descargar todos los ficheros.

Una vez hecho esto, se debe crear el entorno de trabajo. Para ello, simplemente se ejecuta desde la terminal de *Linux* el comando `'conda env create -file env.txt'`, desde el directorio en el que se encuentre el archivo `'env.txt'`.

Los datos no vienen incluidos con el código, se deben descargar de *PLA* [23]. He sido muy poco flexible con los formatos que deben tener para simplificar la tarea de programación. Deben estar guardados en el mismo directorio que el archivo `'smooth.py'`, en carpetas que se llamen `'Components/[componente]'`, donde *componente* es `'CMB'`, `'Synchrotron'`, `'Dust'` o `'Freefree'`, según que componente sea. Los ficheros se deben llamar `'map_[frec]'`, donde *[frec]* es la frecuencia de observación. Una vez hecho esto, crea dos carpetas que se llamen `'CMB'` y `'observations'` en el directorio de partida y ejecuta las funciones `'save.CMB'`, `'make.simulation'` y `'do.mean'` (dejando esta última para el final). Con esto ya se tienen los mapas de trabajo.

Tras esto, se debe crear la plantilla para hacer la separación por regiones. Para ello, simplemente ejecuta la función `'template'` de `'template.py'`, indicando las latitudes a las que quieras dividir el mapa.

La estructura de guardado de los datos de salida planteada es la siguiente: crea una carpeta *output* en el directorio en el que se encuentra la carpeta `'Program'`. En ella, crea una carpeta donde se van a almacenar los datos en formato matricial, `'raw_data'`, y otra para la simulación, `'simulation_[num]'`, donde *[num]* es un entero para identificar la simulación. Para simplificar las cosas, usa `'simulation_1'` para el resultado dividido en regiones, y `'simulation'` para el aplicado al cielo completo.

Con esto, en `'main.py'`, que se encuentra en `'Program/Separation'`, `'prepare_files'` coge los datos previamente preparados y los guarda en formato matricial en `'output/raw_data'` como `'data_raw.fits'`, y *separate* es la separación en sí.

Si has seguido este formato de carpetas sugerido, las pruebas de `'Program/tests'` son inmediatas. Simplemente ejecuta ambos programas y se obtienen los resultados finales.