Quantum aspects originated by Gravitation

from Cosmology to Astrophysics

By

ADRIAN DEL RIO VEGA

VNIVERSITAT DE VALÈNCIA

A dissertation submitted to the University of Valencia, in accordance with the requirements of the "Programa de Doctorat en Física", in order to obtain the degree of DOCTOR OF PHILOSOPHY.

JUNE 2018

Advisors:

Prof. Jose Navarro Salas.
Dr. Ivan Agullo Rodenas.

Abstract

The study of quantum fields propagating in classical, curved and dynamical spacetimes offers a first approach to assess the consequences of the quantum theory when gravitational phenomena are not negligible. This is an important question that must be addressed when an intense gravitational field plays a principal role in the dynamics of a physical system, such as during the early universe (cosmological inflation) or in the formation of astrophysical black holes.

One of the most striking features of this subject is perhaps that the calculation of physical observables, even for non-interacting fields, often involves ill-defined quadratic operators of fields, thus requiring a non-trivial and suitable renormalization method. The normal-ordering operation, usually employed in Minkowski spacetime, no longer works here, since additional ultraviolet divergences associated to curvature arise. The standard approach is to subtract the short-distance asymptotic behaviour of the two-point (Green) functions of interest. As a consequence, unexpected results are predicted due to finite remaining terms coming together with renormalization subtractions, and demanded by general covariance. The goal of this thesis is to give new insights following this direction.

In the first part, we analyze the consequences of renormalization of quantum fields on diverse aspects of inflationary cosmology. Issues related to the ultraviolet divergences arising in computation of the angular power spectrum of the CMB in the Sachs-Wolfe regime, or renormalization of the stress-energy tensor of matter (spin 1/2) fields during inflation, preheating, or the big-bang expansion of the universe, are considered. The implications of "hidden" fields present during single-field inflation are also studied and physical implications coming from CMB bounds are discussed.

Renormalization is in fact of crucial importance since it can even lead to the breakdown of well-known classical symmetries and associated conserved Noether charges, yielding what is normally referred to in the literature as quantum anomalies. In the second part of this thesis we present a new and particularly interesting example of this feature in electrodynamics: the classical E-B duality symmetry of Maxwell equations without charges and currents fails to hold at the quantum level if spacetime has curvature and non-trivial dynamics. In fact, our results suggest that a dynamical curved spacetime with significant frame-dragging is able to distinguish between the two (left and right) circular polarization states of the photons. This offers promising physical implications in binary mergers in astrophysics, whose observational window is nowadays open thanks to the recent detections of gravitational waves by the LIGO-Virgo collaboration.

Dedication and acknowledgements

Si estoy escribiendo esta tesis doctoral a día de hoy es gracias al apoyo y oportunidades brindados por diferentes personas.

Estoy particularmente agradecido a mis supervisores doctorales y mentores, Pepe e Iván, por su ayuda y colaboración en la investigación, y el haber estado siempre abiertos y disponibles para discutir cualquier cuestión que me surgiera. Les debo un gran crecimiento científico a través de sus lecciones y constantes discusiones a lo largo de los años. Han pasado en particular ya siete años desde que conocí a Pepe, y le estaré siempre agradecido por la oportunidad al haber aceptado entrenarme y supervisar mi TFG en su día, así como el buen trato, completa libertad para trabajar, y apoyo para realizar todo lo que me propusiese.

Me gustaría agradecer también a las personas que me dieron la oportunidad de visitar sus respectivos centros de investigación y el colaborar con ellos durante el doctorado. Además de Iván Agulló en Louisiana, éstas incluyen a Ruth Durrer y Subodh Patil en Ginebra, y Abhay Ashtekar en Pennsylvania. Me trataron siempre con mucha amabilidad y como si fuera uno más en sus grupos de investigación. Mención especial también a Anton Joe y Javier Olmedo por su generosidad y siempre valiosa e inolvidable ayuda durante las estancias de investigación.

Agradezco también a mis compañeros de doctorado Nicolás y Paco, por colaboración científica fructífera y momentos divertidos a lo largo de estos años de doctorado y durante todos los viajes a congresos y escuelas, así como al resto del departamento de astrofísica, Álex, Jesus, Isabel, Jose, Tommmmmek, Sergio, Vassili, etc, que sin ellos el doctorado no habría sido ni la mitad de divertido de lo que ha sido. Estoy también muy agradecido a Samuel, por todos los momentos durante estos años; por convencerme a estudiar matemáticas, y por todas las discusiones de clase.

Finalmente, pero no por ello menos importante, mención especial a familia y amigos. Mis padres Lola y Enrique, mis primos José y Marta, mis tíos José, Maria José y Juana, y mis ya fallecidos abuelos Lola y Daniel. Gracias por apoyarme siempre. Sé que a mi abuelo siempre le ha hecho mucha ilusión que me dedicara a estudiar ciencias y matemáticas, pues fue lo que a él le hubiera gustado hacer, pero como a tantas otras personas de la España de la época, no tuvo la más remota oportunidad para ello. Sé también que le hace mucha ilusión a mi padre, que aunque las opciones ya no eran tan desfavorables y pudo estudiar Física, las oportunidades para luego dedicarse a la investigación y el doctorado eran prácticamente nulas. Yo sí puedo decir que he tenido esas oportunidades, y espero haberlas exprimido al máximo. Es por ello que les dedico la tesis a ellos. También me gustaría agradecer a mis amigos de toda la vida, Dani, Soler, Boscà, Rico, Julio, Miguel, y Nico, por todos los momentos divertidos y por estar siempre ahí; y además a mi novia Maque y su familia por todo el apoyo dado, entusiasmo, y sobretodo paciencia.

Agradezco finalmente al Ministerio de Educación de España por haberme concedido la FPU y financiación para las tres estancias de investigación en el extranjero, así como a las trabajadoras del servei d'investigació de la universitat por disponibilidad y paciencia siempre que tenía dudas acerca de todos los trámites involucrados.

AUTHOR'S DECLARATION

I declare that the work presented in this thesis was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes.

The contents of Chapters 2 and 3 and part of 5 are based on the following papers. Credit copyright American Physical Society, World Scientific, and IOP Science.

- "*Gravity and handedness of photons*", I. Agullo, A. del Rio, J. Navarro-Salas, Int. J. Mod. Phys. D, Vol. 26 (2017) 1742001. [First Prize in the 2017 international Essay Competition of the Gravity Research Foundation].
- "*Adiabatic regularization with Yukawa interaction*", A. del Rio, A. Ferreiro, J. Navarro-Salas, and F. Torrenti, Phys. Rev. D 95, 105003 (2017).
- "*Electromagnetic duality anomaly in curved spacetimes*" I. Agullo, A. del Rio, J. Navarro-Salas, Phys. Rev. Lett 118, 111301 (2017).
- "*On the renormalization of ultraviolet divergences in the inflationary angular power spectrum*", A. del Rio, J. Navarro-Salas, Journal of Physics: Conference Series 600 (2015) 012023
- "*Equivalence of adiabatic and DeWitt-Schwinger renormalization schemes*", A. del Rio and J. Navarro-Salas, Phys. Rev. D 91 (2015) 064031.
- "*Renormalized stress-energy tensor for spin-1/2 fields in expanding universes*", A. del Rio, J. Navarro-Salas, and F. Torrenti, Phys Rev D 90 (2014) 084017.
- "*Space-time correlators of perturbations in slow-roll de Sitter inflation*", A. del Rio and J. Navarro-Salas, Phys. Rev. D 89 (2014) 084037.

Some contents of Chapter 4, written in collaboration with R. Durrer and S. Patil [63], and some of Chapter 5, written in collaboration with I. Agullo and J. Navarro-Salas [12], are planned to be sent to some peer-review journals for publication as well, and thus copyrights and convenient permissions shall be transfered to the journal, eventually. It shall follow the respective regulation once the work is published. The text presented here should be understood only as a dissertation submitted to the University of Valencia as required to obtain the degree of Doctor of Philosophy in Physics.

Except where indicated by specific reference in the text, this is the candidate's own work, done in collaboration with, and/or with the assistance of, the candidate's supervisors and collaborators. Any views expressed in the thesis are those of the author.

Valencia, May 4, 2018.

Adrian del Rio Vega.

RESUMEN EN CASTELLANO.

Contexto y motivación

La teoría de la relatividad general es una descripción geométrica del espacio-tiempo y de la gravitación formulada por A. Einstein en 1915 [136, 184]. En esta teoría el espacio-tiempo se describe mediante una variedad diferenciable, dotada de un tensor métrico g_{ab} de signatura Lorentziana para proporcionar una noción física de distancia y causalidad. Lo que caracteriza a este teoría es que el espacio-tiempo se considera una entidad dinámica: la métrica no es fija sino que, por el contrario, puede evolucionar de acuerdo a las ecuaciones diferenciales de Einstein, $G_{ab}(g) = 8\pi G T_{ab}$, donde $G_{ab}(g)$ es el tensor de Einstein construido a partir de la métrica, el cual captura la información geométrica, y T_{ab} describe el contenido energético en el espacio-tiempo. Partículas prueba se propagan en este escenario curvo siguiendo geodésicas, y la interacción gravitatoria se explica de esta manera como una manifestación de curvatura del espacio-tiempo.

Se trata de la teoría satisfactoria más simple que tenemos para interpretar todos y cada uno de los fenómenos gravitatorios observados a día de hoy, dando cuenta de predicciones cuantitativas que han sido comprobadas con gran precisión por muchos y diversos experimentos a lo largo de estos 100 años de historia [191]. Sin embargo, es también bien conocido que la teoría conduce a su propio fracaso al predecir, bajo ciertas circunstancias, la formación de las llamadas singularidades espacio-temporales [113]. Aunque algunas conjeturas proponen la imposibilidad de un observador de "ver" estas singularidades, como por ejemplo la censura cósmica de Penrose [162], está ampliamente aceptado que dichas singularidades están realmente señalando los límites de validez de la teoría. Esto, junto al hecho de que las otras interacciones fundamentales en la Naturaleza son de origen cuántico, sugiere que la relatividad general probablemente emerge como un límite clásico, de bajas energías, de una teoría cuántica de la gravedad.

Aunque varias ideas se han desarrollado para construir tal teoría, la falta total de observaciones a día de hoy que proporcionen información al respecto impone retos muy serios para distinguir cada una de estas teorías, o incluso falsear cualquiera de ellas. Uno entonces se ve limitado a seguir principios teóricos, prejuicios y argumentos, para construir la teoría, con la esperanza de que futuros experimentos comprueben las predicciones asociadas a ella. Un análisis detallado de las diferentes vías pensadas con tal fin está fuera del alcance de este texto, y referimos al lector a consultar [163] para una visión general de la situación.

A falta de una dirección o idea clara y definitiva de lo que una teoría cuántica de gravedad debería ser, ¿podríamos, sin embargo, sacar alguna conclusión acerca de la influencia de gravedad sobre fenómenos cuánticos? Recordemos que en los primeros años de la teoría cuántica numerosos cálculos se desarrollaron asumiendo la interacción de materia, descrita cuánticamente, con un entorno electromagnético clásico, dando lugar a resultados que en última instancia estaban de acuerdo, con alto grado de precisión, con las predicciones de la teoría electromagnética cuántica completa. Se espera que una situación similar ocurra para gravedad, en la cual se pueda estudiar la propagación de materia o radiación, descrita cuánticamente, en un entorno gravitacional clásico. Esta es el problema tratado por la teoría cuántica de campos en espacio-tiempos curvos [48, 94, 157, 185], y proporciona un primer paso para abordar las cuestiones relacionadas con gravedad cuántica. La escala de Planck establece el límite más allá del cual una teoría cuántica de gravedad completa debe ser tenida en cuenta. Se espera que los efectos cuánticos de la gravitación puedan ser despreciados siempre y cuando las distancias involucradas sean mucho mayores que la distancia de Planck, y dado que la longitud de Planck es minúscula, esto parece proporcionar mucho margen de validez a la teoría.

Este camino proporciona una visión semiclásica, menos ambiciosa que una teoría cuántica completa, pero observacional y físicamente más accesible, como por ejemplo la cosmología primordial ha demostrado. En efecto, el trabajo pionero desarrollado por L. Parker en los sesenta [147–150], el cual dio lugar al descubrimiento de producción de partículas por universos en expansión, ha resultado ser fundamental para explicar las inhomogeneidades iniciales en la densidad de energía de materia al inicio de la expansión del big bang. Dichas irregularidades primordiales, en última instancia, y a través del colapso gravitatorio, llevaron a la formación de la estructura a gran escala en el universo que observamos hoy día [72, 77, 189]. Se consideran el resultado de la amplificación de las fluctuaciones cuánticas de un campo escalar (el inflaton) durante la violenta expansión del espacio-tiempo que tuvo lugar durante el universo temprano. Por otro lado, esta teoría es también de gran importancia desde un punto de vista conceptual. En concreto, el descubrimiento por Hawking en 1975 de la emisión térmica de partículas durante un colapso gravitatorio, usando estas técnicas [84, 111], es una piedra angular en física fundamental moderna. Aunque esta radiación no ha sido observada todavía, es una idea muy convincente y ampliamente reconocida porque establece una conexión muy sólida entre relatividad general, teoría cuántica de campos, y física estadística o termodinámica. La visión semiclásica es también relevante pues captura algunos fenómenos no perturbativos, tales como el efecto Schwinger [85, 157] y las llamadas anomalías cuánticas [91–93], y uno verdaderamente espera que una teoría cuántica de gravedad completa también acabe dando cuenta de todos estos fenómenos.

Aunque esta teoría todavía está en sus inicios a nivel experimental, nuevas generaciones de misiones en cosmología y astrofísica, tal como PLANCK [6], SDSS [7], LIGO-Virgo [5], satélites y observatorios Multimessenger [2, 3], etc; y experimentos planeados para un futuro como LISA [4], Einstein telescope [1], etc; están jugando un papel cada vez más importante en testear física fundamental hoy día, y podrían dar lugar a avances muy significativos. Desde hace aproximadamente 25 años, la cosmología disfruta de su edad de oro, numerosos experimentos han sido llevados a cabo con éxito midiendo cantidades de interés físico en diversas formas complementarias, al mismo tiempo que alcanzando precisiones del uno por ciento. De hecho, las misiones cosmológicas actuales son capaces de observar fenómenos físicos de tal altas energías en la cual la gravedad es tan intensa que efectos cuánticos de la materia y la radiación son relevantes para explicar los datos. Por otro lado, la reciente detección directa de ondas gravitacionales por la colaboración LIGO-Virgo, originadas en la fusión de sistemas binarios de agujeros negros o estrellas de neutrones, junto al análogo electromagnético observado en todo el espectro por la astronomía multimessenger, proporcionan acceso único al régimen extremadamente intenso y no lineal del campo gravitatorio. Esta posibilidad abre una nueva ventana para testear la relatividad general a muy altas energías en un contexto astrofísico, y asimismo efectos nuevos podrían ser importantes en la descripción de la física subyacente.

Para concluir, tales crecientes éxitos experimentales requieren predicciones teóricas más precisas de la relatividad general y/o de la teoría cuántica. La tesis que presento apunta en esta dirección, ofreciendo algunos resultados e ideas al respecto. El trabajo realizado se resume en las siguientes dos secciones.

Aspectos cuánticos en cosmología primordial

En la primera parte de esta tesis nos preocuparemos de estudiar efectos cuánticos de campos en cosmología. La cosmología es una rama de relatividad general que estudia el universo como un sistema físico en su totalidad, y cuyo principal objetivo consiste en entender su estructura y dinámica [72, 77, 125, 189]. Aunque en sus orígenes la cosmología se soportaba en conjeturas y las observaciones eran escasas, hoy día existe una cantidad ingente de datos y un marco teórico muy desarrollado que ajusta los datos cuantitativamente con mucha precisión. Existe buena evidencia de que la materia en el universo (en forma de galaxias, cúmulos de galaxias, "materia oscura", etc) se distribuye de manera altamente homogénea e isótropa a escalas suficientemente grandes, y que el espacio tiempo se expande aceleradamente. Estos hechos, junto con el descubrimiento del fondo cósmico de radiación de microondas (CMB, por sus siglas en inglés) ha establecido una teoría de que nuestro universo está en expansión y en enfriamiento, y que se originó de una "gran explosión" (o "big bang" en inglés).

La dinámica de la expansión se describe por medio de una familia de soluciones de las ecuaciones de Einstein con constante cosmológica ($G_{ab}(g) + \Lambda g_{ab} = 8\pi G T_{ab}$), que son espacialmente homogéneas e isótropas. Estas soluciones se conocen como métricas de Fridmann-Lemaitre-Robertson-Walker (FLRW). Para foliaciones con superficies espaciales planas, y parametrizadas en coordenadas cartesianas, se escriben como $ds^2 = dt^2 - a(t)^2 \delta_{ij} dx^i dx^j$, y sólo dependen de una función del tiempo cósmico t, llamado factor de escala $a(t)$, que tiene en cuenta la expansión de estas superficies espaciales con el tiempo cósmico. La velocidad de la expansión se le atribuye al parámetro de Hubble $H = \dot{a}/a$. Para entender la historia del universo uno debe determinar entones la dependencia funcional del factor de escala con el tiempo. La relatividad general conecta esto con el contenido energético T_{ab} del universo; en otras palabras, el problema se reduce a saber cómo la densidad de energía evoluciona con el tiempo cósmico. Pero resulta que esto es realmente una cuestión muy complicada dado que debemos saber los diferentes tipos de especies (materia, radiación, "energía oscura", etc) contribuyendo a ello, cada una de las cuales reescala de manera diferente con el tiempo.

Por otro lado, el CMB es un fondo observado de radiación electromagnética, prácticamente homogéneo e isótropo, que permea todo el universo. Muestra un espectro de frecuencias de cuerpo negro prácticamente perfecto a una temperatura $T_0 = 2.725\ K$. Este fondo de radiación primordial es una predicción del modelo cosmológico estándar del big bang y constituye una imagen muy valiosa del universo temprano. Justo después de que se produjera la "gran explosión", cuando el universo era mucho más caliente y denso que ahora, no había átomos neutros o núcleos formados. Los fotones y electrones en particular estaban en continua interacción manteniendo un equilibro térmico. La alta tasa de interacciones en tal escenario aseguraba que cualquier átomo o núcleo producido sería inmediatamente destruído por un fotón de alta energía. Conforme el universo se expandió y enfrió por debajo de las energía de enlace atómico, los elementos neutros empezaron a formarse. En particular, ls electrones y protones se unieron para dar lugar a átomos de hidrógeno neutros y estables. De esta manera, hubo un tiempo en la evolución del universo (normalmente llamado "superficie de última interacción") en el cual los fotones pararon de interaccionar con la materia, y empezaron a propagarse libremente desde entonces hasta nuestros días, preservando sus propiedades. Estos son los fotones que observamos hoy día y constituyen el CMB. Las minúsculas fluctuaciones de temperatura respecto de una homogeneidad e isotropía perfectas, que se observan al medir los fotones en diferentes direcciones en el cielo, tienen diferentes orígenes físicos, y proporcionan una de las fuentes de información más importantes que tenemos sobre la evolución del universo.

La teoría del big bang, sin embargo, no está libre de inconvenientes. Durante los primeros instantes parece necesitar de un periodo, conocido como inflación [43, 128, 189], en la cual una expansión violenta del espacio-tiempo habría tenido lugar. Este periodo se introduce ad hoc en la teoría del big bang, y se asume que un ajuste suave con la siguiente evolución debe ser posible. Matemáticamente, este universo se modeliza usando el conocido espacio-tiempo de Sitter, cuya métrica puede ser expresada como un caso particular de FLRW con una función exponencial como factor de escala, $a(t) \sim \exp(Ht)$. La importancia del paradigma inflacionario reside en que proporciona un mecanismo sólido y muy convincente de explicar aquellas pequeñas inhomogeneidades ("perturbaciones escalares de densidad", en lenguaje de teoría de perturbaciones cosmológicas lineales [43, 77]) que, tras la evolución cósmica, acaban originando las fluctuaciones tanto de temperatura del CMB como de la distribución de materia a gran escala en el universo.

Y es aquí cuando la teoría cuántica juega un papel fundamental en gravedad. Estas inhomogeneidades primordiales puede ser generadas de manera natural durante la inflación cósmica como resultado de la amplificación de fluctuaciones cuánticas. Como ya hemos avanzado en la subsección anterior, este mecanismo se entiende gracias al proceso cuántico, espontáneo o estimulado, de creación de partículas por la dinámica gravitacional de un universo en expansión, descubierto por L. Parker en los sesenta [147]. Observaciones detalladas en los últimos años indican que las fluctuaciones de temperatura del CMB se ajustan con alto grado de precisión a una distribución gaussiana con un espectro de fourier que es casi invariante de escala [166]. Las observaciones son compatibles, de manera no trivial, con las predicciones de inflación, incluso con la implementación mínima basada en un sólo campo escalar con un potencial de "slow-roll" conduciendo la expansión inflacionaria. La zoología de modelos [133] que predicen los mismos valores para observables inflacionarios disponibles hoy día es considerablemente grande, desafortunadamente, y ello motiva la búsqueda de características adicionales de inflación que puedan distinguir entre todos estos modelos. La generación de ondas gravitacionales ("perturbaciones tensoriales") durante la rápida expansión es también una predicción de inflación, que todavía no ha sido detectada. Algunos experimentos, tal como PLANCK, están buscando señales de gravitones primordiales en la polarización magnética de los fotones del CMB. La detección de estas ondas sería un espaldarazo definitivo a la teoría de inflación, y proporcionarían una fuente valiosa de información del universo temprano para averiguar las incógnitas del modelo inflacionario.

Tal y como bien se conoce en teoría cuántica de campos, el cálculo de observables físicos normalmente requiere tomar valores esperados de operadores compuestos, lo cual resulta no estar bien definido, ya que presentan divergencias ultravioleta (UV) que necesitan ser tratadas adecuadamente. En espacio-tiempos curvos, la cuestión se hace más delicada todavía, incluso para campos libres. La operación convencional de hacer "normal-ordering", usualmente empleada en Minkowski para campos libres, no funciona aquí dado que surgen varias divergencias UV adicionales asociadas a la curvatura del espacio tiempo. La cuestión fue resuelta en los setenta y un resumen de todo ello se puede encontrar en [48]. La idea consiste en sustraer, a la cantidad de interés, el comportamiento asintótico a pequeñas distancias de la función de dos puntos apropiada [94]. Tal y como se encontró en [53] por primera vez, las sustracciones pueden ser reabsorbidas en última instancia en las constantes de acoplo de la teoría, la constante cosmológica Λ, y la constante de Newton G, dando lugar al final a un marco teórico auto consistente.

El principal observable cosmológico que apoya la teoría de inflación es el llamado espectro de potencias angular, y, como veremos en más detalle, su cálculo involucra tomar el valor esperado de ciertos operadores de campo. Es natural preguntarse entonces si este cálculo requiere renormalización. Una cuestión similar fue planteada para el espectro de potencias, la versión de Fourier de la función de dos puntos de perturbaciones primordiales, y el problema se abordó en el pasado por medio del método de renormalización adiabático [16–19, 146]. Estas ideas llevaron, sin embargo, a un cierto debate en la literatura [41, 78, 89, 132, 168], con algunas personas argumentando en contra del uso de renormalización en este contexto. En el capítulo 2 mostraremos el papel de las divergencias UV que aparecen en el espectro de potencias angular y sugeriremos la renormalización correspondiente desde una perspectiva espacio-temporal (no recurriendo al espacio de Fourier). Las ideas presentadas en este capítulo están basadas en dos artículos [65, 67]. Usaremos la función de dos puntos de perturbaciones primordiales generadas durante la inflación para derivar una expresión analítica de los coeficientes multipolares C_ℓ en el régimen de Sachs-Wolfe. Analizamos el correspondiente comportamiento UV y enfatizaremos el hecho de que el resultado estándar actual en la literatura es realmente equivalente a haber hecho renormalización de la función de dos puntos a "0 orden adiabático". Luego argumentaremos que renormalización a "segundo orden adiabático" resulta ser más apropiado desde un punto de vista físico. Esto podría alterar significativamente las predicciones para C_ℓ, sin llegar a romper la invariancia de escala.

Después del periodo de inflación, y previo a la subsiguiente expansión de la teoría del big bang, se asume que existió una etapa intermedia conocida como "recalentamiento", en la cual toda la materia presente hoy día en el universo observado se habría originado. Existen muy pocos datos disponibles que den información de esta etapa, pero se piensa que el campo inflatón, que conduce la expansión inflacionaria, es responsable de crear del vacío todas las partículas del modelo estándar a través de acoplos tipo Yukawa durante este periodo [123, 124]. Con ello, resulta necesario encontrar un procedimiento sistemático de calcular observables físicos renormalizados propios de esta etapa.

Con esta motivación en mente, en el capítulo 3 extendemos el método de renormalización adiabático para calcular el tensor energía-momento de campos de Dirac. Lo calcularemos primero para el caso de un campo libre, y luego para el caso en que el campo tenga un acoplo de Yukawa con un campo escalar de fondo. El formalismo adiabático es una ténica muy conocida y ampliamente usada para campos cuánticos en espacio-tiempos homogéneos basados en una expansión asintótica UV de los modos de fourier del campo, desarrollado por Parker y Fulling [95, 96, 152]. Originalmente, este método fue llevado a cabo solamente para campos Klein-Gordon en universos en expansión. La extensión de la expansión adiabática para modos de campos fermiónicos es sutil y hasta hace poco no se había desarrollado en la literatura. Este método permite calcular observables físicos sistemáticamente (como densidad de energía renormalizada, momento, o el número de partículas creadas) de interés en espacio-tiempos en expansión, como los que normalmente tienen lugar en cosmología. Los resultados de este proyecto se publicaron en tres artículos [64, 66, 68]. El trabajo fue hecho en colaboración con gente de Valencia (A. Ferreiro y J. Navarro-Salas) y de Madrid (F. Torrenti).

En el capítulo 4 vamos más allá y consideramos no sólo los efectos cuánticos del campo que genera la inflación sino el impacto de los efectos cuánticos de otros campos "espectadores". En concreto, estudiamos la posibilidad de extraer cotas sobre el número total de campos en el universo a partir de cotas experimentales del CMB en el espectro de potencias tensoriales. Campos espectadores (que no interaccionan con el inflaton) presentes durante inflación pueden afectar observables del CMB a través de sus fluctuaciones cuánticas, amplificadas por la dinámica gravitatoria: son capaces de inducir una dependencia logarítima en las funciones de correlación de perturbaciones de curvatura (lo que en inglés se conoce un "running"). En este trabajo hemos considerado el efecto de un gran número de tales grados de libertad en los observables inflacionarios, y mostraremos que se puede extraer cotas sobre el contenido de campos espectadores del universo a través de cotas del índice espectral tensorial (asumiendo que el espectro tensorial llegue un día a medirse). El proyecto involucró trabajar intensamente con teoría de perturbaciones lineales en cosmología, inflación, física del CMB, cálculo de funciones de correlación, y diagramas de Feynman dentro del llamado formalismo "in-in". Este proyecto se originó durante una estancia de investigación en el grupo de cosmología y astrofísica de la Universidad de Ginebra en 2016, bajo la supervisión y colaboración de Ruth Durrer y Subodh P. Patil. En las próximas semanas enviaremos el trabajo para publicar.

Aspectos cuánticos en astrofísica

La astrofísica es otro campo de la gravitación que estudia el origen, formación, evolución, y propiedades de las diferentes estructuras individuales en el universo, tales como estrellas, galaxias, cúmulos de galaxias, agujeros negros, rayos cósmicos, materia oscura, etc. Paralelamente a la inflación, los intensos fenómenos gravitacionales que tienen lugar en típicas situaciones de astrofísica podrían sevir para testear la validez de la relativiad general y / o teoría cuántica de campos en espacio-tiempos curvos. En particular, la formación de agujeros negros, típicamente produciéndose como consecuencia del colapso gravitatorio de una estrella, resulta ser un escenario interesante posible para comprobar las implicaciones de campos cuánticos propagándose alrededor, como el popular efecto Hawking [111]. Otro escenario significativo, que es particularmente popular hoy día, es aquel de un sistema binario de objetos compactos que acaban fusionándose para dar lugar a un agujero negro, y que durante el proceso liberan una cantidad ingente de energía en forma de ondas gravitacionales y electromagnéticas. Recientemente, el obsevatorio de ondas gravitacionales por interferómetro láser (LIGO, por sus siglas en inglés) en EEUU, conjuntamente con el interferómetro Virgo en Europa, han detectado por primera vez, y con métodos directos, las ondas gravitatorias emitidas durante la fusión de dos agujeros negros [5]. De hecho, la colaboración ha detectado más eventos de fusión de binarias, y no solo resultado de la colisión de dos agujeros neros, sino también de la fusión de dos estrellas de neutrones. El homólogo electromagnético esperado para este último caso también fue detectado, casi simultáneamente, y en todo el espectro electromagnético, por diversos telescopios y observatorios [2, 3], marcando con ello un avance significativo para la astronomía multi-messenger también. Las observaciones resultantes proporcionan acceso único a las propiedades del espacio-tiempo en el régimen de gravedad intensa, altas velocidades, y confirman las predicciones de la relatividad general para la dinámica no lineal de agujeros negros altamente perturbados.

La segunda parte de esta tesis trata enteramente con la simetría de dualidad eléctrico-magnética de la teoría de Maxwell sin fuentes a nivel cuántico, asumiendo que el campo electromagnético se propaga en un espacio-tiempo clásico, curvo y dinámico. Este tema será cubierto en detalle en el capítulo 5. Aunque originalmente estuvimos interesados en posibles aplicaciones en cosmología, tal como discutiremos las implicaciones físicas de este fenómeno resultarán ser mayoritariamente relevantes en astrofísica, principalmente porque hay disponibles escenarios gravitatorios con cierto grado de vorticidad o "frame-dragging". El proyecto fue iniciado en una estancia de investigación en Louisiana State University en 2015, y continuó en la Universitat de València, bajo la supervisión y estrecha colaboración del Dr. Ivan Agullo y Prof. Jose Navarro Salas.

Clásicamente, la acción de Maxwell para la electrodinámica en ausencia de cargas y corrientes eléctricas es invariante bajo una rotación de dualidad de los campos eléctricos y magnéticos, incluso cuando el campo electromagnético está inmerso en un espacio-tiempo clásico, curvo y dinámico. Asociada a esta simetría existe una carga Noether conservada que en Minkowski mide la diferencia neta entre fotones polarizados circularmente a izquierdas y derechas (y con ello resulta estar relacionado con el parámetro V de Stokes). La simetría de dualidad garantiza que ello sea una constante del movimiento en ausencia de fuentes electromagnéticas. En el capítulo 5 mostraremos, sin embargo, que cuando el campo se cuantiza, esta simetría deja de ser válida, proporcionando lo que se conoce como una anomalía en la teoría. Puede ser interpretada como el análogo de spin 1 de la llamada anomalía quiral para fermiones, y como ésta, podría proporcionar implicaciones físicas interesantes. La estrategia que seguiremos consistirá en simular la teoría fermiónica tanto como sea posible, en particular reescribiendo la acción clásica como aquella de Dirac.

El origen de esta anomalía puede ser entendida como una consecuencia de la renormalización: cuando uno trata con una simetría clásica, el teorema de Noether nos dice que la divergencia de la corriente resulta ser proporcional a las ecuaciones de movimiento, con coeficientes que son los campos en sí. Esto da lugar a un operador cuadrático en los campos, y como tal necesita ser renormalizado adecuadamente. Las sustracciones de renormalización no necesariamente respetan las ecuaciones de movimiento de los campos dado que la restricción principal es que sean covariantes, y de aquí la aparición de la anomalía. Esto es precisamente lo que ocurre con otras anomalías. Mostraremos también, complementariamente, que la anomalía puede ser entendida siguiendo la interpretación de Fujikawa. Las amplitudes de transición entre dos estados diferentes del campo electromagnético pueden ser calculadas usando lo que se conoce como la integral de camino. Esto requiere sumar sobre todos los caminos posibles en el espacio de fase entre ambos estados inicial y final, cada uno de ellos pesado por la exponencial de la acción. Una transformación de dualidad es una transformación canónica en el espacio de fase, por lo que debe dejar la amplitud de transición (o acción cuántica efectiva a primer orden) invariante. Se sabe entonces que la acción permanece invariante también, debido a la simetría clásica. Sin embargo, a diferencia de la teoría clásica, además de la acción hay una ingrediente adicional, que aparece dentro de la integral de camino, que debe ser analizado: la medida de la integral podría cambiar en un jacobiano no trivial, en efecto.

Ambos cálculos proporcionan el mismo resultado y confirman la existencia de esta anomalía cuántica. Independientemente de la interpretación dada, la conclusión es que contribuciones "off-shell" (fuera de las ecuaciones de movimiento) del campo electromagnético anulan la simetría clásica.

El trabajo presentado involucró tratar y estudiar una amplia gamma de temas, tales como dinámica hamiltoniana, spinores, teoría de campos, cuantización, teorías gauge, integrales de camino, anomalías cuánticas, formalismo 3+1 en relatividad general, teoría de la renormalización en espacios curvos, análisis geométrico, etc. Los primeros resultados fueron publicados en Physical Review Letters [13]. Es de destacar también que un ensayo [14], que propone y enfatiza potenciales implicaciones físicas de este efecto cuántico en astrofísica, ganó el primer premio en la prestigiosa competición de ensayos de la Gravity Research Foundation en 2017. Una continuación del trabajo con muchos más detalles está siendo preparada para publicación [12].

Desde un punto de física más físico, esta anomalía sugiere que un campo gravitatorio distingue ambos grados de libertad radiativos de los fotones (es decir, fotones polarizados circularmente tanto a derecha como a izquierda) induciendo un cambio en su estado de polarización a través de fluctuaciones cuánticas, posiblemente durante la propagación. En particular, hemos aprendido recientemente que la anomalía se manifiesta si el campo gravitatorio admite la emisión de ondas gravitatorias. Consecuentemente, el resultado podría ser observacionalmente relevante en espacio-tiempos dinámicos con cierto grado de vorticidad o "frame-dragging", tal como el colapso de una estrella de neutrones a un agujero negro de Kerr, o fusión de sistemas de binarias compactas en astrofísica. Estos últimos son considerados una de las fuentes de emisión de ondas gravitatorias más importantes y se piensa que constituyen algunos de los eventos más violentos que tienen lugar en el universo. Los métodos analíticos empleados en estos casos son muy limitados, y como consecuencia se colabora con expertos en relatividad numérica (N. Sanchis Gual, V. Mewes, J.A. Font) para estimar el efecto en situaciones astrofísicas típicas. Se han conseguido algunos resultados, pero el proyecto está todavía en desarrollo. Aquí comentaremos sólo las ideas principales desde un punto de vista cualitativo. Esta colaboración ha permitido al autor familiarizarse con aspectos clave en el campo de la relatividad numérica [20] (condiciones gauge, el problema de los datos iniciales, escalares de Weyl para extraer información en la emisión de ondas gravitatorias, etc).

Como avanzábamos en el párrafo anterior, el tema de emisión de ondas gravitatorias está fuertemente conectado con la manifestación de la anomalía en la dualidad electromagnética. Esto es realmente una conclusión altamente no trivial, y para llegar a ella hemos empleado técnicas en relatividad general que versan sobre análisis asintótico. Estos métodos son bien conocidos para la descripción de espacio-tiempos asintóticamente planos, un marco teórico que resulta muy conveniente para el estudio de fuentes astrofísicas aisladas, tales como agujeros negros en colapso o dinámicos, y la propagación de radiación emitida por ellos hacia el futuro nulo (los métodos son realmente válidos para cualquier campo radiativo, no sólo gravitacional). Temas como el "infinito nulo conforme" y "compactificación a la Penrose", expansiones "peeling-off"; base nula de Newman-Penrose; dinámica hamiltoniana covariante; simetrías asintóticas BMS; cargas infrarrojas "suaves" y transformaciones gauge "grandes"; el efecto memoria, etc; son clave en estos asuntos. Fueron estudiados por el autor mientras visitaba el Instituto para la Gravitación y el Cosmos en 2017 en una estancia de investigación supervisada por Abhay Ashtekar. Aunque este trabajo todavía se encuentra en desarrollo, cierto progreso se ha conseguido y algunos aspectos serán comentados al final de capítulo 5.

Finalmente, damos por finalizada la tesis resumiendo los aspectos principales de los diferentes capítulos del texto y describiendo pronósticos futuros de trabajo al respecto. Esto se lleva a cabo en el capítulo 6, que pasamos a resumir a continuación en esta traducción a castellano.

Conclusiones y futuras direcciones de investigación

Esta tesis es el resultado de la investigación llevada a cabo por el autor durante los últimos 5 años en colaboración con sus supervisores y otros investigadores. El tema central es la teoría cuántica de campos que se propagan en espacio-tiempos curvos. Aunque ya han pasado más de 50 años desde el primer artículo cuantitativo al respecto, [147], el tema es todavía de gran importancia hoy día, no sólo porque continua dando nuevos hallazgos en las bases de una supuesta teoría cuántica de la gravedad, sino por que conduce a nuevas implicaciones fenomenológicas que, hoy más que nunca, podrían ser testadas en experimentos. A día de hoy, las misiones cosmológicas han alcanzado tal nivel de sensibilidad que los efectos cuánticos de los campos presentes durante el universo temprano deben ser tenidos en cuenta para la correcta interpretación del análisis de datos en ambos CMB y estructura a gran escala. Por otro lado, la detección de ondas gravitatorias por la colaboración LIGO-Virgo, junto al análogo electromagnético medido por la astronomía multi-messenger, ha abierto una nueva era en astrofísica. No hay duda de que este área de investigación en rápido desarrollo proporcionará una cantidad considerable de datos de fuentes astrofísicas que permitirán un mejor entendimiento no sólo de relatividad general, sino también de las consecuencias de la teoría cuántica de campos alrededor de escenarios tipo agujero negro.

Con esta motivación en mente, en la primera parte de esta tesis hemos trabajado en ciertas cuestiones relacionadas con cosmología primordial. En el capítulo 2 hemos analizado el observable más importante que apoya la teoría inflacionaria, el espectro de potencias angular, usando una perspectiva diferente a la convencional. Reanalizamos esta cantidad desde un punto de vista espacio-temporal y encontramos que el resultado estándar considerado en la literatura es realmente divergente en el límite ultravioleta. La hipótesis usual de tomar el límite de distancias enormes en la función de dos puntos, que acaba llevando a la bien conocida (y experimentalmente comprobada) invariancia de escala del espectro, sirve de regulador natural para estas divergencias UV, y resulta ser equivalente a una sustracción de orden adiabático 0 para la función de dos puntos involucrada. Hemos argumentado que el segundo orden adiabático debería ser más natural desde un punto de vista físico. Al hacer esto uno obtiene un término adicional (también invariante de escala) en el espectro de potencias (véase (2.39) para consultar el resultado final), y sus posibles consecuencias observables fueron discutidas.

El capítulo 3 constituye un texto detallado que versa sobre el cálculo del tensor energía-momento renormalizado de campos de spin 1/2 en universos en expansión. El trabajo se llevó a cabo tanto cuando el campo fermiónico se propaga libremente como cuando interacciona con un campo escalar a traves de un acoplo tipo Yukawa. Expresiones finales de los valores renormalizados pueden ser consultados en (3.72) y (3.73); y (3.157), (3.158), (3.171). Esto permite el estudio de la materia (concretamente, la producción de partículas y energía) no sólo durante el régimen inflacionario, sino también durante el recalentamiento, durante el cual toda la materia conocida en el universo se supone que fue originada a través de las oscilaciones del campo inflatón. Algunos ejemplos de interés cosmológico usando este formalismo fueron discutidos usando aproximaciones analíticas (véase sección 3.4 y apéndice B) y algunos usando métodos numéricos (véase apéndice E). El trabajo que se presenta allana el camino para futuros proyectos relacionados con aplicaciones numéricas.

El estudio cosmológico acaba con el capítulo 4. En modelos de inflación se asume normalmente que el único campo relevante para el cálculo de observables físicos es aquel que conduce la expansión exponencial. En ese capítulo hemos considerado la presencia e influencia de un gran número N de campos escalares ligeros "espectadores", es decir, que no interacciónan con el campo inflatón sino solo se acopla a gravedad. Esto es de interés dado que algunas teorías fundamentales (por ejemplo teoría de cuerdas, supergravedad, etc) generalmente requieren de la existencia de ingentes cantidades de tales campos ligeros por auto-consistencia, y su presencia debería ciertamente tener algún tipo de consecuencia observacional, por ejemplo en observables del CMB. Resulta que contribuyen al "running" en ambos espectros escalar y tensorial. En este trabajo hemos derivado este resultado usando métodos alternativos a los disponibles en la literatura. Hemos discutido además que, aunque normalmente son despreciados en los cálculos involucrados de inflación, su contribución en el espectro tensorial podría ser muy relevante para poner cotas en el posible número N de campos escalares ligeros presentes en el universo. La cota obtenida y la discusión asociada pueden consultarse en (4.58). Los resultados podrían restringir ciertos modelos fenomenológicos que requieren de tales ingentes cantidades de campos ligeros para resolver el problema de la jerarquía en el modelo estándar.

Volviendo ahora a aspectos más fundamentales de la teoría cuántica de campos, en el capítulo 5 hemos tratado con la transformación de dualidad eletrocmagnética de la teoría de Maxwell sin fuentes en espacio-tiempos curvos. Se sabe que clásicamente es una simetría de la acción de Maxwell si no hay cargas y corrientes eléctricas. Sin embargo, motivados por el conocimiento de otras anomalías, se planteó y analizó en detalle la cuestión natural de si esta simetría clásica podría ser extendida a la teoría cuántica. En efecto, tal y como se mostró por primera vez en 1969 [10, 44, 122], esto es un asunto no trivial: las simetrías clásicas pueden fallar en la teoría cuántica debido a contribuciones off-shell procedentes de correcciones cuánticas. En este capítulo hemos calculado el valor esperado en vacío de la divergencia de la corriente Noether, y hemos encontrado que no es nula debido a las sustracciones de renormalización, las cuales son necesarias con el fin de lidiar adecuadamente con las divergencias ultravioletas. Este resultado conduce por tanto a una anomalía en la teoría cuántica. Hemos comprobado además el cálculo siguiendo la interpretación de Fujikawa: la medida de la integral de camino, que determina la amplitud de transición entre estados cuánticos del campo a diferentes tiempos, se transforma no-trivialmente, proporcionando la anomalía. Hemos acabado el capítulo comentando implicaciones físicas potencialmente interesantes en astrofísica.

Respecto a este último capítulo, el trabajo realizado abre varias direcciones de investigación en el futuro, las cuales pasamos a describir a continuación. Nos surgen cuestiones inmediatas: como deberíamos interpretar esta anomalía de dualidad electromagnética desde un punto de vista físico? Significa que una muestra inicial de fotones se puede polarizar si atraviesan un campo gravitatorio fuerte? Más aún, predicen algunas características distintivas las desviaciones de la teoría de la relatividad general? Existen implicaciones observables de todo ello, digamos, en astrofísica? Surge esta anomalía en la radiación gravitatoria también? Planeamos abordar todas estas cuestiones en un futuro cercano trabajando una serie de direcciones de investigación específicadas abajo. Responder dichas preguntas nos ayudará a entender las implicaciones de esta anomalía desde un punto de vista medible, y esto será útil para testear la validez de la relatividad general y la teoría cuántica de campos en espacio-tiempos curvados en un contexto astrofísico.

1. El cálculo concreto de la anomalía de dualidad electromagnética fue desarrollado considerando el valor esperado en vacío de la corriente de dualidad, es decir, hemos calculado la cantidad $\langle 0|\nabla_\mu j^\mu|0\rangle \neq 0$. Este efecto de polarización podría ser interpretado como creación espontánea de partículas (asimétricamente) del vacío debido a la fuerte dinámica gravitacional, similar al celebrado efecto Hawking. Si, en lugar de ello, el estado inicial que describe el campo electromagnético se asume inicialmente con una distribución de fotones, esperaríamos una contribución estimulada a este fenómeno desde el estado cuántico, lo cual debe ser mucho más relevante en cualquier situación práctica de interés cosmológico y astrofísico. Nuestro plan consiste en determinar cuan grande este efecto puede llegar a ser. En otras palabras, necesitamos calcular $\langle\rho\big|\nabla_\mu j^\mu\big|\rho\rangle$ en detalle, donde $\big|\rho\rangle$ representa un estado mixto que describe fotones con diferentes helicidades, energías, etc. Con todo ello, en última instancia tendremos que averiguar cómo esta cantidad podría ser medida por un instrumento cuántico.

Utilizaremos técnicas adicionales de teoría cuántica de campos en espacio-tiempos curvos y relatividad general. El uso de detectores Unruh-DeWitt, en particular, podría ser relevante para entender en mayor profundidad cuál es el impacto de la anomalía sobre los fotones desde un punto de vista físico, dado que estos detectores nos dicen qué es lo que un campo test u observador sería capaz de medir o no. Análisis asintótico en relatividad general será importante también para abordar el estudio de la radiación procedente del pasado nulo infinito y la radiación dispersada hacia el futuro nulo infinito. Por otro lado, métodos usualmente empleados en ciencia de la información cuántica podrían ser muy relevante además para analizar la cuestión de la elección del estado cuántico inicial, por ejemplo por medio de operadores densidad.

2. Una suposición adicional en nuestro cálculo de la anomalía de dualidad electromagnética es el uso de relatividad general convencional para describir la gravitación. Dado que el resultado obtenido es una contribución geométrica pura, podría ser interesante investigar si teorías modificadas de gravedad proporcionan nuevos términos a $\langle 0|\nabla_{\mu j^\mu}|0\rangle$ or $\langle\rho\big|\nabla_{\mu j^\mu}\big|\rho\rangle$. En tal caso, esto podría servir para predecir desviaciones de la relatividad general en entornos gravitacionales intensos, como por ejemplo a través de nuevas características en las ondas gravitatorias emitidas. Diversos caminos pueden emprenderse tanto en teoría de campos como en geometría diferencial. Por ejemplo, si permitimos torsión o no-metricidad en la conexión, además de la curvatura contribuciones adicionales a los valores esperados de arriba podrían aparecer.

3. Las ondas gravitatorias también tienen dos grados de libertad radiativos, en concreto los bien conocidos modos h_+ and $h_\times$, y existe una noción análoga a la simetría de dualidad electromagnética. Consecuentemente, a nivel lineal en la teoría, la noción de una carga que mida el estado de polarización de la radiación debería estar disponible también. Planeamos analizar esto en detalle y estudiar si existe una anomalía en la teoría de gravitones que se propagan en un espacio-tiempo curvo no trivial. Es decir, si el entorno gravitacional distingue (a través de fluctuaciones cuánticas) entre los dos grados de libertad de las ondas gravitatorias. Esto podría ser particularmente interesante en futuras medidas de polarización con ondas gravitacionales.

4. Nuestro plan final sería buscar aspectos distintivos de todo ello en astrofísica, en concreto en física de agujeros negros y emisión de ondas gravitatorias. Trabajar en este área implicará familiarizarse con fuentes astrofísicas, tales como fusión de binarias o agujeros negros dinámicos, así como con métodos analíticos involucrados para extraer de ellos información de interés físico (análisis asistólico, momentos multipolares de agujeros negros, cargas asistólicas, etc). Se abordarán cuestiones como por ejemplo si el efecto crece linearmente con la masa del sistema, o si la polarización o el momento angular llevados por las ondas juega algún papel. Se espera estrecha colaboración con relativistas numéricos para poder realizar simulaciones.

Table of Contents

Page

CHAPTER 1

INTRODUCTION

1.1 Overview and motivation

General relativity is a geometrical description of spacetime and gravitation formulated by A. Einstein in 1915 [136, 184]. In this theory spacetime is described by a smooth manifold, endowed with a Lorentzian metric g_{ab} to provide a physical notion of distance and causality. The key feature is that spacetime is regarded as a dynamical entity: the metric is not fixed but rather it can evolve according to Einstein's hyperbolic differential equations, $G_{ab}(g) = 8\pi G T_{ab}$, where $G_{ab}(g)$ is the Einstein tensor constructed from the metric, capturing geometric information, and T_{ab} describes the energy content of spacetime. Test particles propagate in this curved background by following geodesics, and the gravitational interaction is thus explained as a manifestation of curvature in spacetime.

This is the simplest successful theory we have for interpreting all gravitational phenomena observed so far, accounting for concrete predictions that have been tested to great accuracy by many different experiments during these 100 years of history [191]. However, it is also well known that it leads to its own failure by predicting the formation of spacetime singularities under certain circumstances [113]. Although some conjectures regarding the impossibility of an observer to "see" these singularities are formulated, such as the cosmic censorships [162], it is widely accepted that these features are actually pointing out the limits of validity of the theory. This, together with the fact that the other fundamental interactions in Nature are of quantum origin, suggests that general relativity probably emerges as the classical low-energy limit of a quantum theory of gravity.

Although different attempts have been developed for constructing such a theory, the lack of observations so far raises serious challenges in distinguishing between them, or even to falsify any of them. One is then limited to follow theoretical principles, prejudices and arguments to construct the theory, with the hope that future experiments will check the predictions. A detailed description of the different approaches in this respect is out of the scope of this text, and the reader is referred to [163] for an overview of the situation.

Lacking a clear or definite idea of what a quantum theory of gravity should be, can we still hope to learn something about the influence of gravity on quantum phenomena? Recall that in the old history of the quantum theory, numerous calculations were performed assuming the coupling of quantum matter to a classical electromagnetic background field, yielding results that eventually were in agreement with the full quantum electrodynamics theory to an excellent degree of approximation. We may expect that a analogous situation would occur for gravity, in that the propagation of quantized matter or radiation fields in a classical gravitational background is studied. This is the subject of quantum field theory in curved spacetimes [48, 94, 157, 185], and provides the first step in our way for addressing all these questions. The Planck length indicates the limit beyond which the full quantum gravity theory must be recalled. It is expected that quantum gravitational issues are negligible when physical distances involved are much larger than the Planck length, and since Planck scale is so tiny this seems to provide much scope for the framework.

This approach yields a semiclassical picture, less ambitious than a full quantum theory, but observationally and physically more accesible, as for instance primordial cosmology has shown. Indeed, the pioneering work developed by L. Parker in the 60's [147–150], resulting in the discovery of particle production by expanding universes, has been proven to be fundamental to explain the initial seeds in matter energy density at the beginning of the big bang expansion that eventually, through gravitational clump, led to the formation of the large scale structure of the universe that we observe today [72, 77, 189]. These primordial inhomogeneties in density are regarded as the result of the amplification of quantum fluctuations of a scalar field during the violent expansion of spacetime that took place during the early universe. On the other hand, this approach is also important from a conceptual point of view. Namely, the discovery by Hawking in 1975 of thermal emission of particles by gravitational collapse using these techniques [84, 111] is a cornerstone in modern fundamental physics. Although this radiation has not been observed so far, it is very compelling and widely recognized because it establishes a strong connection between general relativity, quantum field theory, and statistical physics or thermodynamics. The semiclassical picture is also very relevant since it captures some non-perturbative phenomena, such as the Schwinger effect [85, 157] and the so-called quantum anomalies [91–93], and one expects that a full quantum theory of gravity should account for all these phenomena as well.

Although this theory is still experimentally very limited, new generations of observational missions in cosmology and astrophysics, such as PLANCK [6], SDSS [7], LIGO-Virgo [5], Multimessenger detections [2, 3], etc; and future planned experiments such as LISA [4], Einstein telescope [1], etc; are playing an increasingly important role in testing fundamental physics nowadays and could lead to significant advances. For around 25 years now cosmology enjoys its golden age, numerous experiments have been carried out measuring quantities of physical interest in diverse complementary ways, while at the same time reaching sensitivities of one-hundred percent. In fact, currently cosmological missions are able to test energy scales in which gravity is so strong that quantum effects of matter and radiation become relevant. On the other hand, the recent direct detection of gravitational waves by the LIGO-Virgo collaboration, emitted from the merger of black hole and neutron star binaries, together with the electromagnetic counterpart observed in the whole spectrum by the multimessenger astronomy community, provide unique access to the strong field, non-linear regime of gravity. This availability opens a new window to test general relativity at very high energies in an astrophysical context, and again some novel effects might be important in describing some of the underlying physics.

Overall, such increasing experimental achievements demand for more accurate predictions coming from general relativity and/or quantum theory. The presented thesis points towards this direction, offering some results and insights concerning this. They are briefly summarized in the following two sections.

Quantum aspects in primordial cosmology

In the first part of this thesis we shall be concerned on quantum effects of fields in cosmology. Cosmology is a branch of general relativity that studies the Universe as a whole physical system, and whose main goal consists in understanding its structure and dynamics. [72, 77, 125, 189]. Although originally cosmology relied on certain assumptions and observations were very sparse, today there is a huge amount of data and a well-developed theoretical framework that coincides quantitatively with it. Today there is good evidence that matter in the universe (in the form of galaxies, clusters of galaxies, "dark matter", etc) appears highly homogeneous and isotropic at sufficiently large scales, and that spacetime expands acceleratedly. This, together with the discovery of the cosmic microwave background (CMB), has settled down the well-known "big bang" theory of an expanding and cooling universe that originated from an initial singularity, in which all matter and energy was concentrated.

The dynamics of the expansion is described by a family of solutions to Einstein's equations with a cosmological constant, $G_{ab}(g)+\Lambda g_{ab}=8\pi G T_{ab}$, satisfying spatial homogeneity and isotropy. These are called Fridmann-Lemaitre-Robertson-Walker (FLRW) metrics. For spatially flat slices they are written as $ds^2=dt^2-a(t)^2\delta_{ij}dx^i dx^j$ using cartesian coordinates, and only depend on a function of cosmic time t, called scale factor $a(t)$, which takes into account the expansion of spacetime. The velocity of the expansion is attributed to the Hubble parameter, $H=\frac{\dot{a}}{a}$. To understand the history of the Universe one must determine then the functional dependence of the scale factor with cosmic time. General relativity relates this with the energy content T_{ab} filling the universe, i.e. the problem is reduced to know how the energy density develops with comic time. But this is actually a complicated issue since we must know the different kinds of species (matter, radiation, "dark energy", etc) contributing to it, each of which scale distinctly with time.

On the other hand, the CMB is an observed background of electromagnetic radiation almost perfectly homogeneous and isotropic filling the whole universe. It shows an almost perfect black body frequency spectrum at a temperature of $T_0=2.725\ K$. This background of primordial radiation is a prediction of the standard big bang cosmological model and constitutes a highly valuable picture of the early universe. Right after the "big bang" singularity, because the temperature and density of the universe were much higher, neutral atoms or bound nuclei could not form. Photons, electrons and the rest of charged particles were in continuous interaction mantaining a thermal equilibrium. The high rate of interactions in such an environment guaranteed that any nucleus or atom produced would be instantly annihilated by a high energy photon. As the universe expands, its temperature decreases until it gets below characteristic binding energies, allowing this way light elements to form. In particular, electrons and protons join to form neutral, and stable, hidrogen atoms. This way, there was a time in the evolution of the universe (normally called "last scattering surface") in which photons stopped interacting with matter and started propagating freely since then until our days, preserving their properties. These are the photons that we observe nowadays and constitute the CMB. The small temperature fluctuations from a perfect homogeneity and isotropy, that are found by measuring the photons at different directions in the sky, have different physical origins, and provide one of the most major references that we have about the evolution of the universe.

The "big bang" theory, though, is not lack of drawbacks. It appears to need a period during the early universe, called inflation [43, 128, 189], in which a violent expansion of spacetime is supposed to take place. This is introduced ad hoc and a smooth match with the subsequent big bang expansion is expected. Mathematically this universe is modeled using the well-known de Sitter spacetime, whose metric can be expressed as a particular case of FLRW with an exponential function as the scale factor: $a(t) \sim \exp(Ht)$. The importance of the inflationary paradigm is that it provides a solid and compelling way to explain the initial little inhomogeneities ("scalar density perturbations", in the language of linear cosmological perturbation theory [43, 77]) that, after the cosmic evolution, are observed both through the temperature fluctuations of the CMB photons, and at the large scale distribution of matter in the universe.

And here is when the quantum theory plays a fundamental role in gravity. These primordial inhomogeneties can be naturally generated during a cosmic inflation period as a result of the amplification of quantum fluctuations. As already advanced in the previous subsection, this mechanism is understood thanks to the spontaneous or stimulated quantum process of particle creation by the gravitational dynamics of the expanding universe, discovered by L. Parker in the 60's [147]. Detailed observations in the last years indicate that temperature fluctuations of the CMB fit with high degree of precision to a Gaussian distribution with a Fourier spectrum which is almost scale invariant [166]. The observations are compatible, in a non-trivial way, with the predictions of inflation, even with the minimal implementation based in a single scalar field with a "slow-roll" potential driving the expansion. The zoo of models [133] predicting current inflationary observables is considerably large, though, and this motivates the search for additional signatures of inflation that may distinguish between all models. The generation of gravitational waves ("tensorial perturbations") during the rapid expansion is also a prediction of inflation, that still has not been detected. Some experiments such as Planck are seeking for signals of primordial gravitons in the magnetic polarization of CMB photons. The detection of these waves is a smoking gun of the theory of inflation, and will provide a rich source of information from the very early universe to figure out the unknowns of the inflationary model.

As it is well-known from quantum field theory, the calculation of physical observables normally requires taking expectation values of quadratic operators, ill-defined quantities that present UV divergences and need to be treated suitably. In curved spacetime the issue becomes delicated even for free fields. The conventional normal-ordering operation usually employed in Minkowski spacetime for free fields no longer works here since several UV divergences associated with the curvature of the spacetime background arise. The issue was solved in the 1970's, and nicely summarized in [48]. The idea consists in subtracting the asymptotic short-distance behaviour of the appropriate two-point function to the quantity of interest [94]. As first noted in [53], the subtractions can eventually be reabsorbed in the coupling constants of the theory, the cosmological constant Λ and the Newton constant G, yielding at the end a self-consistent framework.

The main cosmological observable supporting the theory of inflation is the angular power spectrum, and, as we shall see in more detail, its calculation involves taking the vacuum expectation value of some field operators. It is natural then to ask whether this computation requires renormalization. A similar question was posed for the power spectrum, the Fourier space version of the two-point function of primordial perturbations, and several attempts were carried out in the past by means of the so-called adiabatic renormalization scheme [16–19, 146]. These ideas leaded, though, to some debate in the literature [41, 78, 89, 132, 168], with some people arguing against the use of renormalization here. In Chapter 2 we show the role of ultraviolet divergences arising in the angular power spectrum and suggest the corresponding renormalization, from a spacetime perspective. The ideas presented in this chapter are based on two papers [65, 67]. We employ the two-point function of primordial perturbations generated during inflation to derive an analytic expression for the multipole coefficients C_ℓ in the Sachs-Wolfe regime. We analyze then the corresponding ultraviolet behavior and stress the fact that the current standard result in the literature is actually equivalent to a renormalization of the two-point function at zeroth adiabatic order. Then we argue that renormalization at second adiabatic order is more appropriate from a physical point of view. This modifies the predictions for C_ℓ, while keeping scale invariance.

After the period of inflation, and previous to the standard "big bang" expansion of the universe, it is widely understood that an intermediate stage known as "reheating" takes place, in which all the matter present today in our universe is supposed to have originated. There is little amount of information available concerning this stage, but it is thought that the scalar field driving the inflationary expansion is responsible for creating, out from the vacuum, all the particles from the Standard Model through Yukawa coupling during this period [123, 124]. Thus, it appears to be necessary to find a systematic way of calculating renormalized physical observables originating from this stage.

With this motivation in mind, in Chapter 3 we extend the adiabatic renormalization method to calculate the stress-energy tensor of Dirac fields. This is done first for a free fermion field, and then for the case in which this field has a Yukawa coupling with a background scalar field. The adiabatic formalism is a well-known and widely used renormalization technique for quantum fields in homogeneous spacetimes based on an asymptotic UV expansion of the field Fourier modes, developed by Parker and Fulling [95, 96, 152]. Originally it was developed only for Klein-Gordon fields in expanding universes. The extension of the adiabatic expansion of modes to fermions is subtle and until very recently it was lack in the literature. This method allows to compute physical observables sistematically (like renormalized energy density, momentum, or the created amount of particles) of interest in expanding spacetimes, such as the ones normally taking place in cosmology. Three papers of this project were published in Phys. Rev. D [64, 66, 68]. This work was done in collaboration with other people, both in Valencia (A. Ferreiro and J. Navarro-Salas) and Madrid (F. Torrenti).

In chapter 4 we go further and consider not only the quantum effects of the field(s) driving inflation, but rather the impact of quantum effects coming from others, "spectator" fields. Namely, we study the possibility of extracting tensor bounds on the hidden field content of the Universe. Spectator (free) fields present during single-field inflation can affect CMB observables through quantum fluctuations: they induce logarithmic running in correlation functions of curvature perturbations. In this work we considered the effect of a large number of such field degrees of freedom on inflationary observables, and showed that one can extract bounds on the hidden field content of the universe through bounds on the tensor running of the spectral index. The project involves dealing with linear perturbation theory in cosmology, inflation, CMB physics, calculation of correlation functions, and Feynman diagrammatics within the so-called in-in formalism. This project was originated during a research stay at the Cosmology and Astrophysics group in Geneva in 2016, under the supervision and collaboration of Prof. Ruth Durrer, and Dr. Subodh Patil. A paper concerning these results is being prepared for communication [63].

Quantum aspects in astrophysics

Astrophysics is another field of gravitation that studies the origin, formation, evolution, and properties of the different individual structures in the Universe, such as stars, galaxies, clusters of galaxies, black holes, cosmic rays, dark matter, etc. Parallel to inflation, the intense gravitational phenomena that take place in some typical situations in astrophysics could help to test the validity of general relativity and / or quantum field theory in curved spacetimes. In particular, the formation of black holes, typically arising as a consequence of the gravitational collapse of a star, turns out to be a possible interesting scenario to check the implications of quantum fields propagating nearby, such as the celebrated Hawking effect [111]. Another significant scenario, which is particularly popular nowadays, is that of binary systems of compact objects that end up coalescing to give a black hole, and in the process release a huge amount of energy in the form of either gravitational and electromagnetic waves. Recently, the Laser Interferometer Gravitational-Wave Observatory (LIGO) in the USA, jointly with the Virgo interferometer in Europe, have detected for the very first time, and with direct methods, the gravitational waves emitted during the merger of two black holes [5]. The collaboration has in fact detected more events supporting this, and not only from the collision of two black holes, but also from the merger of two neutron stars. The expected electromagnetic counterpart for this latter case was also detected in the whole electromagnetic spectrum by several telescopes and observatories, thereby representing a valuable breakthrough for multi-messenger astronomy as well. The observations provide deep insight in the behaviour of spacetime in the regime of intense gravity, high velocities, and confirm the predictions of general relativity for the non-linear dynamics of two black holes [5].

The second part of this thesis deals entirely with the electric-magnetic duality transformation of source-free Maxwell theory in the quantum regime, assuming the electromagnetic field to propagate in a classical and dynamical curved spacetime. This topic will be covered in detail in chapter 5. Although originally we were interested in possible applications to cosmology, as we shall argue physical implications of this phenomenon will turn out to be mostly relevant in astrophysics, mainly because gravitational scenarios with vorticity or frame-dragging are available. The project was initiated in a research stay at Louisiana State University in 2015, and continued at the University of Valencia, under the supervision and close collaboration of both Dr. Agullo and Prof. Navarro-Salas, respectively.

Classically, Maxwell action for electrodynamics in absence of electric charges and currents is invariant under a duality rotation of the electric and magnetic fields, even when the electromagnetic field is immersed in a classical and dynamical curved spacetime. Associated to this symmetry there is a conserved Noether charge (the V-stokes parameter) which measures the net difference among right- and left-handed circularly polarized photons. The duality symmetry states that this is a constant of motion in the absence of electromagnetic sources. We will find, though, that when the electromagnetic field is quantized, this symmetry fails to hold, providing a so-called anomaly in the theory. It can be seen as the spin-1 analog of the so-called chiral anomaly for fermions, and as the latter, it may provide interesting physical implications. The strategy that we shall follow will consist in mimicking the fermion theory as much as possible, in particular in rewriting the classical action as that of a Dirac one.

The origin of this anomaly can be understood in terms of renormalization: when one deals with a classical symmetry, Noether's Theorem tells that the divergence of the current turns out to be proportional to the equations of motion, with coefficients being the fields themselves. This results in a quadratic operator field and as such needs to be renormalized properly. Renormalization subtractions need not respect the equations of motion of the fields since the primary restriction is that they are covariant, and thus the emergence of an anomaly. This is precisely what happens in other anomalies. We also show, complementary, that the anomaly can be understood following Fujikawa's interpretation. Transition amplitudes between two different states of the electromagnetic field are calculated using the path integral. This requires summing over all paths in phase space, each of which weighted with the exponentiated action. A duality transformation amounts to a canonical transformation in phase space, so must leave the whole transition amplitude (or one-loop effective action) invariant. One knows then that the action remains invariant, because of the classical symmetry. However, unlike in the classical theory, besides the action there is an additional ingredient that must be analyzed and appears inside the path integral: the measure may change in a non-trivial jacobian, indeed. Both calculations yield the same result and confirm the existence of this quantum anomaly. Regardless of the interpretation, the conclusion is that quantum off-shell contributions of the electromagnetic field spoil the classical symmetry.

The project involves dealing with a wide variety of topics, such as hamiltonian dynamics, spinors, field theory, quantization, gauge theories, path integrals, anomalies, 3+1 formalism of gravity, renormalization theory in curved spacetime, geometric analysis, etc. The first results were published in Physical Review Letters [13]. It is also worth to remark that an essay [14], proposing and emphasizing potential physical implications of this quantum effect in astrophysics, won the First award in the prestigious Gravity Research Foundation Essay Competition in 2017. A follow-up paper with more details is being prepared for publication [12].

From a more physical point of view, this anomaly is suggesting that a gravitational background distinguishes both radiative degrees of freedom of photons (i.e., right- and left- handed circularly polarized photons), by inducing a change in their polarization state through quantum fluctuations, possibly along the propagation. In particular, we have recently learnt that the anomaly manifests if gravitational wave emission takes place in the background. Consequently, the result could be observationally relevant in dynamical spacetimes with a certain degree of vorticity or frame-dragging, such as the collapse of a neutron star into a Kerr black hole, or mergers of compact binary systems in astrophysics. These are considered one of the most important sources for gravitational-wave emission and are thought to be at the origin of some of the most violent events in the Universe. Analitycal approaches are very limited in this direction, and as a consequence I collaborate with experts in numerical relativity (N. Sanchis-Gual, V. Mewes, J.A. Font) to estimate the effect in typical astrophysical situations. Some results have been obtained but the project is still ongoing. Here we shall comment only the main ideas from a qualitatively point of view. This interaction in particular allowed the author to get familiarized with key issues in the field of numerical relativity [20] (gauge conditions, initial data, Weyl scalars to infer gravitational waves emission, etc).

As advanced in the previous paragraph, the issue of gravitational wave emission is strongly connected to the manifestation of this anomaly. This is actually a non-trivial conclusion and to arrive at this we employ techniques in general relativity dealing with asymptotic analysis. These are well-known methods to deal with asymptotically flat spacetimes, a framework that is nice to describe isolated astrophysical sources, such as dynamical or collapsing black holes, and the propagation of radiation emitted by them to future null infinity (the methods are actually valid for any radiative field, not only gravitational). Conformal null infinity and Penrose compactification, peeling-off expansions, Newman-Penrose null basis, covariant hamiltonian dynamics, BMS asymptotic symmetries, soft (infrared) charges and "large" gauge transformations, "memory" effect, etc, are key topics related to this framework. These were studied by the author while visiting the Institute for Gravitation and the Cosmos in 2017, under the supervision of Prof. Abhay Ashtekar. Although this work is still ongoing, some progress and aspects are discussed by the end of Chapter 5 using the above techniques to figure out that the electromagnetic duality anomaly is related to polarized gravitational wave emission. In particular, it manifests if emission of gravitational waves takes place.

We finally end the thesis by summarizing the main aspects throughout the text and describing future prospects of our work. This is done in Chapter 6.

1.2 Methodology

The methods employed in this thesis are essentially theoretical (focused on calculating observables of physical interest, consult the bibliography, proposing models to explain phenomenology, predicting results given some theories or assumptions, relating ideas to produce new results, increase knowledge in both physics and mathematics, clarifying or giving solutions to questions, etc) and/or computational in some cases (use of advanced calculus software such as Mathematica). It also required the interaction and collaboration with researchers in other fields of expertise (for instance to run advanced numerical simulations in relativity). In general, we use tools from diverse areas of physics (such as classical mechanics, classical theory of fields, quantum field theory, electrodynamics, gauge theories, general relativity, cosmology, and astrophysics) as well as of mathematics (linear algebra, real and complex calculous, classical and modern differential geometry, functional analysis, and geometric analysis). Specific methods have already been commented along the introductory text.

1.3 Training

Apart from research, as part of the Ph.D thesis training I tried to improve my knowledge and education both in physics and mathematics. This is fundamental to gain clarity, insight, perspective, rigor, new and powerful methodology, etc. In particular, I tried to cover a considerable number of books for developing the thesis (see the bibliography) and, in order to increase my background in theoretical physics, I attended several international summer schools abroad whose lectures ranged from general relativity to quantum field theory. Here there is a list of them:

- "*School on Gravitational waves for Cosmology and Astrophysics*", Centro de Ciencias Pedro Pascual, Benasque (Spain), May 28^{th} - June 10^{th}, 2017.
- "*Foundations and new methods in Theoretical Physics*", Wolfersdorf (Germany), September 4^{th} - 16^{th}, 2016.
- "*Geometric aspects of General Relativity*", University of Montpellier (France), September 28^{th} - October 1^{st}, 2015.
- "*Group theory in Particle Physics*", Institut de Fisica Corpuscular, Valencia (Spain), June 2^{nd} - 12^{th}, 2015.

- "*100 years of General Relativity, from theory to experiment and back* ", Israel Institute for Advanced Studies, Jerusalem (Israel), december 29^{th} 2014 - 8^{th} January 2015.
- "*Mathematical Relativity*", Erwin Schrodinger Institute for Mathematical Physics, Vienna (Austria), July 28th - August 1th, 2014
- "*Asymptotic analysis in General Relativity*", Joseph Fourier Institute, Grenoble (France), June 16^{th} - July 4^{th}, 2014.

On the other hand, I also took several lectures in geometry and analysis from the degree of Mathematics, both at the University of Valencia and at the Spanish National University of Distance Education. The courses were: Topology, Differential Geometry of Curves and Surfaces, Hilbert Spaces and Fourier Analysis, Algebraic Topology, Functional Analysis, Differentiable Manifolds and Lie Groups. These lectures provided me a better and solid understanding of the mathematics underlying quantum theory, general relativity, and classical mechanics.

Finally, in order to improve my communication skills, I attended and participated in several conferences all around the world, and gave seminars in different departments. Here there is a list of them:

- "*Electromagnetic duality anomaly in curved spacetime and possible applications to astrophysics*", seminar given at IGC Pennsylvania State University, September 18^{th} 2017.
- "*CMB bounds on the hidden universe*", 12^{th} Iberian Cosmology meeting, Valencia 10-12 April 2017, Spain.
- "*Electromagnetic duality anomaly in curved spacetime*", Oviedo V posgraduate meeting on theoretical physics, Oviedo 17-18 November 2016, Spain.
- "*Electromagnetic duality in curved spacetime*", 21st International Conference on General Relativity and Gravitation, New York 11-15 July 2016, USA.
- "*Electromagnetic duality anomaly*", seminar given at University of Geneva, March 11^{th} 2016.
- "*Adiabatic regularization for spin 1/2 fields and the renormalized stress-energy tensor*", 14^{th} Marcel Grossmann Meeting, Rome 12^{th} - 18^{th} July, 2015.

- ”*The role of renormalization in curved space-time*”, seminar given at IFIC University of Valencia, 20^{th} November, 2014.
- ”*Space-time analysis of primordial perturbations during slow-roll inflation*”, Spanish relativity meeting, 1^{st} - 5^{th} September, 2014.
- ”*Evidence for quantum effects in gravity: from spontaneous particle production to BICEP2 discovery*”, seminar given at IFIC University of Valencia, 29^{th} April, 2014.

CHAPTER 2

Analysis of correlators in inflation and the UV behaviour of the angular power spectrum

Two-point correlators $\langle\mathscr{R}(x)\mathscr{R}(x')\rangle$ and self-correlators $\langle\mathscr{R}^2(x)\rangle$ of cosmological primordial perturbations $\mathscr{R}(x)$ in quasi-de Sitter spacetime backgrounds are considered. For large distance separation between points we find that this correlation function exhibits nearly scale invariance, in agreement to what appears in the power spectrum, its well-known counterpart in Fourier space. Self-correlators, which quantify the amplitude of perturbations at a spacetime point, are ill-defined and need standard renormalization.

We employ then the two-point function of primordial perturbations generated during inflation to derive an explicit expression for the angular power spectrum C_ℓ in the Sachs-Wolfe regime. We study this without the commonly assumed large distance separation, and find that it actually diverges. Although the two objects are of different nature, the renormalized value $\langle\mathscr{R}^2(x)\rangle_{ren}$ can be successfully reproduced using the short-distance behavior of the two-point function $\langle\mathscr{R}(x)\mathscr{R}(x')\rangle$ by a point-splitting version of adiabatic regularization. Motivated by the above, we study then the deformation or "renormalization" of two-point correlators in order to make them smoothly match the renormalized self-correlators at coincidence.

We analyze the ultraviolet behavior of the angular power spectrum and stress the fact that the accepted result in the literature is actually equivalent to a renormalization of the two-point function at zero adiabatic order. We argue that subtractions up to second adiabatic order seems to be more appropriate from a physical point of view. The corresponding angular power spectrum is evaluated in the Sachs-Wolfe regime of low multipoles. Scale invariance is maintained, but the amplitude of C_ℓ could change in a nontrivial way.

The content of this chapter is based on papers [65, 67]

2.1 Motivation for considering renormalization

Before turning our attention to explicit calculations of correlators, let us discuss why one should be interested in the very short-distance behaviour of the two point function, and its potential implications in physical observables.

As commented in the introduction, the theory of quantum fields interacting with gravity [48, 157], applied to the very early and rapidly expanding universe, explains well the pattern of temperature anisotropies of the cosmic microwave background (CMB) as well as the large scale structure (LSS) of the universe. The calculation of observables in a curved space-time is, however, not trivial: new ultraviolet (UV) divergences arise in the computation of local vacuum expectation values, and these infinities can not be removed by standard methods in Minkowski space-time. Specific methods to define regularization and renormalization in expanding universes have been constructed to account for the new UV divergences sourced by curved backgrounds [48, 157] (for more recent works regarding Dirac fields, see [66, 68, 101, 102, 126, 127]).

Let φ represent a generic free field living in an homogeneous and isotropic *Fridmann-Lemaitre-Robertson-Walker* (FLRW) spacetime, with line element $ds^2 = dt^2 - a^2(t)d\vec{x}^2$. The field φ will later describe scalar (or tensor) perturbations during inflation. In the quantum theory the free field operator is most generally studied by its expansion in Fourier k-modes $\varphi_k(t)$,

$$\varphi(t,\vec{x}) = \int d^3k \left[A_{\vec{k}}\varphi_k(t) + A^\dagger_{-\vec{k}}\varphi^*_k(t) \right] e^{i\vec{k}\vec{x}} \ , \tag{2.1}$$

where $A_{\vec{k}}$ and $A^\dagger_{\vec{k}}$ are creation and annihilation operators, such that $A_{\vec{k}}|0\rangle = 0$, with $|0\rangle$ the vacuum state, and satisfy canonical commutation relations, $[A_{\vec{k}}, A_{\vec{k}'}] = 0$,$[A^\dagger_{\vec{k}}, A^\dagger_{\vec{k}'}] = 0$ and $[A_{\vec{k}}, A^\dagger_{\vec{k}'}] = \delta^3(\vec{k} - \vec{k}\,')$. A basic object in quantum field theory is the two-point function. For future purposes, we will consider it at equal times $t = t'$

$$\langle \varphi(t,\vec{x})\varphi(t,\vec{x}')\rangle = \int d^3k |\varphi_k(t)|^2 e^{i\vec{k}(\vec{x}-\vec{x}')} \ . \tag{2.2}$$

We can also construct local physical observables of physical interest. For instance, the local quantum fluctuations of φ can be quantified by the mean square fluctuation in the vacuum state

$$\langle \varphi^2(t,\vec{x})\rangle = \int d^3k |\varphi_k(t)|^2 \ . \tag{2.3}$$

In fact, it is quite common in cosmology to refer the quantity $\Delta^2_\varphi(k,t) \equiv 4\pi k^3 |\varphi_k(t)|^2$ as the power spectrum.

A proper definition of the physical power spectrum in inflationary cosmology is not free of subtleties, as first pointed out in [146], and subsequent studied in [16–19, 153]. In momentum-space, and for a single mode k, the power spectrum $\Delta^2_\varphi(k,t)$ is well defined. However, the formal variance $\langle \varphi^2(\vec{x},t)\rangle$, which is a sum in all modes, diverges in the ultraviolet. There is no doubt that the self-correlator needs renormalization when it is used to quantify the amplitude of quantum perturbations at a single space-time point, as in (2.3). However, the two-point function (2.2) does not need a priori renormalization when used to quantify physical observables involving correlations. One could argue that the two-point function has a well-defined definition in the distributional sense and there is no mathematical need for any regularization [41]. However, as we shall shortly see, the spacetime perspective gives another insight to this question: as the spatial points get close together, the two-point function will grow without bound and produce divergences in some physical observables [65, 67] (for an advance see expression (2.6) together with (2.7)). In [16–19, 146, 153] it was argued that the physical power spectrum should be defined in terms of renormalized quantities. The challenging of this proposal for inflationary cosmology and quantum gravity has been recently stressed in [192]. The purpose of this work is to reanalyze these issues, specially from a spacetime viewpoint.

As we have said, (2.3) is UV divergent and needs to be renormalized according to standard rules. It turns out that, although the two-point function (2.2) and (2.3) are mathematical different objects, the renormalized value of (2.3) can be calculated in terms of the two point function by means of a point-splitting version of adiabatic regularization:

$$\langle \varphi^2(t,\vec{x})\rangle = \lim_{\vec{x}\to\vec{x}'} [\langle \varphi(t,\vec{x})\varphi(t,\vec{x}')\rangle - {}^{(N)}G_{Ad}((t,\vec{x}),(t,\vec{x}'))] \ , \tag{2.4}$$

where ${}^{(N)}G_{Ad}((t,\vec{x}),(t,\vec{x}'))$ is the Nth order adiabatic subtraction term. It is defined via the adiabatic regularization method, or, equivalently, using the DeWitt-Schwinger scheme (for more details see [66].) To properly cancel the UV divergences in our 4-dimensional spacetime the second adiabatic order $N = 2$ is the right one for the mean square fluctuation $\langle \varphi^2(t,\vec{x})\rangle$. It must be noted that the proper adiabatic order of the subtraction term depends on the particular physical quantity to evaluate. For instance, the computation of the renormalized expectation value of the stress-energy tensor $\langle T_{\mu\nu}(t,\vec{x})\rangle$ needs subtraction up to the fourth adiabatic order, using $[\langle\varphi(t,\vec{x})\varphi(t,\vec{x}')\rangle - {}^{(4)}G_{Ad}((t,\vec{x}),(t,\vec{x}'))]$ instead of $\langle\varphi(t,\vec{x})\varphi(t,\vec{x}')\rangle$, and taking the coincident limit.

Apart from these fundamental objects, we can also be interested in integrated quantities from the two-point function, like

$$\langle\varphi_{\vec{p}}\varphi_{\vec{p}'}\rangle = \int d^3\vec{x}d^3\vec{x}' e^{i(\vec{p}\cdot\vec{x}+\vec{p}'\cdot\vec{x}')}\langle\varphi(t,\vec{x})\varphi(t,\vec{x}')\rangle = |\varphi_k(t)|^2\delta^3(\vec{p}+\vec{p}')\ , \tag{2.5}$$

or, assuming rotational invariance, we can also construct

$$\begin{aligned} C_\ell^{\varphi\varphi} &= \frac{1}{4\pi}\int d^2\hat{n}\, d^2\hat{n}' P_\ell(\hat{n}\cdot\hat{n}')\langle\varphi(t,\vec{x})\varphi(t,\vec{x}')\rangle \\ &= 2\pi\int_{-1}^{1} d\cos\theta P_\ell(\cos\theta)\langle\varphi(t,\vec{x})\varphi(t,\vec{x}')\rangle = 16\pi^2\int_0^\infty \frac{dk}{k}|\varphi_k(t)|^2 j_\ell^2(k|\vec{x}|)\ , \end{aligned} \tag{2.6}$$

where P_ℓ are the Legendre Polynomials, and $\cos\theta = \vec{n}\cdot\vec{n}'$ is the angle formed by the two directions $\vec{n} = \vec{x}/|\vec{x}|$ and $\vec{n}' = \vec{x}'/|\vec{x}|$. The physical motivation for considering this formula will become clear in the third subsection of next section. For completeness, we have also added the equivalent and familiar expression in momentum space. It involves the spherical Bessel functions j_ℓ.

Altough the former integral (2.5) is UV finite and nothing else is needed, the latter has an UV divergence. The underlying reason is that the Legendre polynomials do not decay to zero as $\theta \to 0$. A detailed inspection reveals that the divergence is of $N = 0$ adiabatic order. This conclusion can be deduced from the short-distance (asymptotic/adiabatic UV) expansion ($\theta \to 0$)

$$\langle\varphi(t,\vec{x})\varphi(t,\vec{x}')\rangle \sim \frac{1}{1-\cos\theta} - \frac{(\frac{1}{6}-\xi)R}{2}\log(1-\cos\theta) + ... \tag{2.7}$$

and $P_\ell(\cos\theta) \sim 1 + o(\theta)$. Therefore, while (2.5) should be kept unaltered, expression (2.6) must be modified, perhaps according to the renormalization prescription. One should then replace $\langle\varphi(t,\vec{x})\varphi(t,\vec{x}')\rangle$ by $[\langle\varphi(t,\vec{x})\varphi(t,\vec{x}')\rangle - {}^{(0)}G_{Ad}((t,\vec{x}),(t,\vec{x}'))]$ in expression (2.6) to guarantee the UV finiteness of the integral.

However, we should remark that the appropriate choice of the adiabatic subtraction order n depends on the physically relevant object. In CMB cosmology the direct physical observables are temperature correlations, which are linked to the space-time two-point function $\langle\varphi(t,\vec{x})\varphi(t,\vec{x}')\rangle$. For instance, in the Sachs-Wolfe regime we have $\langle\Delta T(\vec{n})\Delta T(\vec{n}')\rangle_{SW} = \frac{T_0^2}{25}\langle\varphi(t,\vec{x})\varphi(t,\vec{x}')\rangle$, where $\varphi(t,\vec{x}) \equiv \mathscr{R}(t,\vec{x})$ is the comoving curvature perturbation. These temperature correlations are the ones upon which other observables like C_ℓ^{TT} are constructed

$$C_\ell^{TT} = \int_{-1}^{1} d\cos\theta P_\ell(\cos\theta)\langle \Delta T(\vec{n})\Delta T(\vec{n}')\rangle \ . \tag{2.8}$$

Since physical correlations are direct observables that can be measured in an experiment, it seems natural to demand that they must always be finite, even at coincidence $\vec{x} = \vec{x}'$. To achieve that, one should then relate $\langle \Delta T(\vec{n})\Delta T(\vec{n}')\rangle$ with the quantity $[\langle\varphi(t,\vec{x})\varphi(t,\vec{x}')\rangle - {}^{(2)}G_{Ad}((t,\vec{x}),(t,\vec{x}'))]$ to ensure UV finiteness at coincidence. Therefore, we argue that the actual related multipole coefficients should be constructed with second order adiabatic subtractions

$$C_\ell^{N=2} = 2\pi \int_{-1}^{1} d\cos\theta P_\ell(\cos\theta)[\langle\varphi(t,\vec{x})\varphi(t,\vec{x}')\rangle - {}^{(2)}G_{Ad}((t,\vec{x}),(t,\vec{x}'))] \ . \tag{2.9}$$

2.2 Spacetime correlators at very large distances

Correlator of tensorial perturbations in a slow-roll scenario

Tensorial perturbations (gravitational waves) generated during inflation can be mathematically described by two independent, massless scalar fields propagating in the unperturbed quasi de Sitter spacetime background. These two scalar fields represent the two independent polarizations of the tensorial fluctuation modes $\mathscr{D}_{ij}$ arising in the inflationary universe. In cartesian coordinates this can be expressed as $ds^2 = dt^2 - a^2(t)(\delta_{ij} + \mathscr{D}_{ij})dx^i dx^j$.

Indeed, working in this metric and set of coordinates, and expanding the fluctuating fields $\mathscr{D}_{ij}$ in Fourier modes $\mathscr{D}_k(t)e_{ij}e^{i\vec{k}\vec{x}}$, where e_{ij} is a constant polarization tensor obeying the conditions $e_{ij} = e_{ji}, e_{ii} = 0$ and $k_i e_{ij} = 0$, the theory of linear cosmological perturbations (see [43, 77, 189]) eventually leads to the equation $\ddot{\mathscr{D}}_k + 3H\dot{\mathscr{D}}_k + \frac{k^2}{a^2}\mathscr{D}_k = 0$, with $k \equiv |\vec{k}|$ and $H = \dot{a}/a$. This is the Klein-Gordon equation in Fourier space for a massless minimally coupled scalar field in a FLRW spacetime. The conditions for the polarization tensor imply that theperturbation field $\mathscr{D}_{ij}$ can be decomposed into two polarization states described by a couple of massless scalar fields $\mathscr{D}_{ij} = \mathscr{D}_+ e^+_{ij} + \mathscr{D}_\times e^\times_{ij}$, where $e^s_{ij} e^{s'}_{ij} = 2\delta_{ss'}$ ($s = +, \times$ stands for the two independent polarizations), both obeying the above wave equation (see, for instance, [189]). For simplicity we omit the subindex $+$ or $\times$.

In the slow-roll approximation of inflation one assumes that the Hubble parameter $H(t)$ changes very gradually, and the change is parametrized by a slow-roll parameter $\epsilon \equiv -\dot{H}/H^2 \ll 1$. Within this approximation it is possible to solve the wave equation in a closed form in terms of the conformal time $\eta \equiv \int dt/a(t)$. Taking into account that $(1-\epsilon)\eta = -\frac{1}{aH}$, the wave equation for $\mathscr{D}_k$ turns out to be of the form

$$\frac{d^2\mathscr{D}_k}{d\eta^2} - \frac{2}{\eta(1-\epsilon)}\frac{d\mathscr{D}_k}{d\eta} + k^2\mathscr{D}_k = 0 \,. \tag{2.10}$$

Treating now the parameter ϵ as a constant, the general solution is a linear combination of Hankel functions [105]. One can univocally fix the solution with the requirement of recovering, for $\epsilon \to 0$, the Bunch-Davies vacuum [171]. The properly normalized solutions for the modes are

$$\mathscr{D}_k(t) = \frac{\sqrt{16\pi G}}{\sqrt{2(2\pi)^3 a^3}}(-\eta\, a\, \pi/2)^{1/2} H_\nu^{(1)}(-k\eta) \,, \tag{2.11}$$

where G is the Newton constant and the index of the Hankel function is exactly $\nu = \frac{3}{2} + \frac{\epsilon}{1-\epsilon}$. Having the explicit form of the modes, we can now compute the two-point function. At equal times $t = t'$ we find (the calculation follows closely to [171])

$$\langle \mathscr{D}(t,\vec{x})\mathscr{D}(t,\vec{x}')\rangle \;=\; \frac{G}{\pi a^2\eta^2}\Gamma\left(\frac{3}{2}+\nu\right)\Gamma\left(\frac{3}{2}-\nu\right){}_2F_1\left(\frac{3}{2}+\nu, \frac{3}{2}-\nu; 2; 1-\frac{(\Delta x)^2}{4\eta^2}\right) \,, \tag{2.12}$$

where $\Delta x \equiv |\vec{x}-\vec{x}'|$. For $\nu = 3/2$ ($\epsilon = 0$) we have the unavoidable infrared divergence of the Bunch-Davies vacuum [23, 24, 90, 182].

For large separations, $a\Delta x \gg H^{-1}$, one obtains

$$\langle \mathscr{D}(t,\vec{x})\mathscr{D}(t,\vec{x}')\rangle \sim \frac{4G\Gamma(3/2-\nu)}{\pi^{3/2}}\frac{\Gamma(\nu)}{a^2\eta^2}\left(\frac{\Delta x}{-\eta}\right)^{2(\nu-3/2)} \,. \tag{2.13}$$

One can immediately observe that the amplitude above is nearly scale invariant, i.e. it almost does not depend on the separation distance Δx. This is consistent with well-known results working in the Fourier domain. Moreover, the term $(-\eta)^{1-2\nu}/a^2$ is time independent, which allows us to evaluate it at the most convenient time. In fact, the correlator can be rewritten as

$$\langle \mathscr{D}(t,\vec{x})\mathscr{D}(t,\vec{x}')\rangle \sim -\frac{16\pi G}{2\epsilon}\left(\frac{H(t_{\Delta x})}{2\pi}\right)^2 \,, \tag{2.14}$$

where the time $t_{\Delta x}$ is defined as $a(t_{\Delta x})\Delta x = H^{-1}(t_{\Delta x})$. Note that there is an implicit slight Δx-dependence on $H(t_{\Delta x})$, given by the one in (2.13).

Correlator of scalar perturbations in a slow-roll scenario

Scalar perturbations originated during inflation can be studied through the gauge-invariant field $\mathscr{R}$ (the comoving curvature perturbation; see, for instance, [189]). For single-field inflation, the modes of the scalar perturbation are given by

$$\mathcal{R}_k(t) = \sqrt{\frac{-\pi\eta}{4(2\pi)^3 z^2}} H_\nu^{(1)}(-\eta k) , \tag{2.15}$$

where now $\nu = 3/2 + (2\epsilon + \delta)/(1-\epsilon)$ and $\delta \equiv \ddot{H}/2H\dot{H}$ is a second slow-roll parameter. Moreover, $z \equiv a\dot{\phi}_0/H$, where $\phi_0(t)$ is the homogeneous part of the inflaton field. These modes determine the vacuum state of scalar perturbations. Such a state can also be regarded as the natural extension of the Bunch-Davies vacuum of de Sitter space. The corresponding two-point function $\langle \mathcal{R}(t,\vec{x}), \mathcal{R}(t,\vec{x}')\rangle$ is given by

$$\langle \mathcal{R}(t,\vec{x})\mathcal{R}(t,\vec{x}')\rangle \quad = \quad \frac{1}{16\pi^2 z^2\eta^2}\Gamma\left(\frac{3}{2}+\nu\right)\Gamma\left(\frac{3}{2}-\nu\right){}_2F_1\left(\frac{3}{2}+\nu,\frac{3}{2}-\nu;2;1-\frac{(\Delta x)^2}{4\eta^2}\right) \tag{2.16}$$

For separations larger than the Hubble radius $a|\vec{x}-\vec{x}'| \gg H^{-1}$ we get

$$\langle \mathcal{R}(t,\vec{x})\mathcal{R}(t,\vec{x}')\rangle \sim \frac{\Gamma(\frac{3}{2}-\nu)}{4\pi^2 z^2\eta^2}\frac{\Gamma(\nu)}{\sqrt{\pi}}\left(\frac{\Delta x}{-\eta}\right)^{2(\nu-3/2)} . \tag{2.17}$$

Again, this is almost scale invariant since $\nu - 3/2 << 1$, as expected from standard results using the Fourier spectrum. This expression can be rewritten, assuming $\nu - 3/2 \equiv (1-n)/2 \approx 0$ (n is the scalar spectral index), as

$$\langle \mathcal{R}(t,\vec{x})\mathcal{R}(t,\vec{x}')\rangle \sim -\frac{4\pi G}{(1-n)\epsilon}\left(\frac{H(t_{\Delta x})}{2\pi}\right)^2 . \tag{2.18}$$

Angular power spectrum. Recovering standard results

Restricting the two-point function of scalar perturbations to points such that $|\vec{x}| = |\vec{x}'|$, we can further obtain $\Delta x^{1-n} = 2^{\frac{1-n}{2}}|\vec{x}|^{(1-n)}(1-\cos\theta)^{(1-n)/2}$ where θ is the angle formed by $\vec{n} = \vec{x}/|\vec{x}|$ and $\vec{n}' = \vec{x}'/|\vec{x}|$. Then, taking $|\vec{x}| = r_L$, where r_L is the comoving radial coordinate of the last scattering surface [189]

$$r_L = H(t_0)^{-1}a(t_0)^{-1}\int_{\frac{1}{1+z_L}}^{1}\frac{dx}{\sqrt{\Omega_\Lambda x^4 + \Omega_M x + \Omega_R}} , \tag{2.19}$$

with the standard cosmological values for z_L, Ω_Λ, Ω_R, and Ω_M [166], the correlator of scalar perturbations for large separations (2.17) shows exactly

$$\langle \mathcal{R}(t,\vec{x})\mathcal{R}(t,\vec{x}')\rangle \quad \sim \quad \frac{4\pi G}{\epsilon}\frac{H^2(1-\epsilon)^2}{16\pi^2}\frac{4\Gamma\left(2-\frac{n}{2}\right)}{\sqrt{\pi}}\Gamma\left(\frac{n-1}{2}\right)2^{\frac{1-n}{2}}\bar{r}_L^{1-n}(1-\cos\theta)^{\frac{1-n}{2}} \tag{2.20}$$

where we have defined the dimensionless quantity $\bar{r}_L(t) \equiv H(1-\epsilon)ar_L$.

This two-point function is related to CMB observations as follows. Let $\Delta T(\hat{n})$ be the temperature fluctuation observed in the CMB for a given direction in the sky $\hat{n} = (\theta, \phi)$, with respect to the (angular) average value of today's temperature T_0, i. e. $\Delta T(\hat{n}) \equiv T(\hat{n}) - T_0$, with $T_0 = \frac{1}{4\pi}\int d^2\hat{n}\, T(\hat{n})$. This magnitude $\Delta T(\hat{n})$ depends on the direction of observation in the sky $\hat{n}$. Since what we observe is a real-valued function which takes values on the spherical 2D surface of the sky, it is convenient to expand it in spherical harmonics $Y_\ell^m(\hat{n})$, which constitutes a complete orthonormal basis: $\Delta T(\hat{n}) = T(\hat{n}) - T_0 = \sum_{\ell m} a_{\ell m} Y_\ell^m(\hat{n})$, where $\ell = 0, 1, 2\ldots$ and the sum goes from $m = -\ell$ to $m = +\ell$. This ℓ is known as the multipole and represents a certain angular scale in the sky by means of the approximation $\theta \approx \pi/\ell$. Since $\Delta T(\hat{n})$ is real, $a^*_{\ell m} = a_{\ell -m}$ holds.

Observational missions measure the angular power spectrum, C_ℓ, defined by

$$\langle a_{\ell m} a^*_{\ell' m'}\rangle = \delta_{\ell\ell'}\delta_{mm'} C_\ell, \tag{2.21}$$

where brackets denote some average process. This formula in turn implies $C_\ell = \langle |a_{\ell m}|^2\rangle$, which shows that C_ℓ are real and definite positive. Now we can calculate the two point function of temperature correlations as

$$\langle \Delta T(\hat{n})\Delta T(\hat{n}')\rangle = \sum_{\ell m} C_\ell Y_\ell^m(\hat{n}) \left(Y_\ell^m\right)^* (\hat{n}') = \sum_\ell C_\ell \frac{2\ell+1}{4\pi} P_\ell(\hat{n}\cdot\hat{n}'), \tag{2.22}$$

where P_ℓ are Legendre polynomials. One can now take an inverse Legendre transformation and obtain a formula for the multipole coefficients C_ℓ in terms of this correlator:

$$C_\ell = \int d^2\hat{n}\, d^2\hat{n}' \frac{P_\ell(\hat{n}\cdot\hat{n}')}{4\pi} \langle \Delta T(\hat{n})\Delta T(\hat{n}')\rangle = 2\pi \int_{-1}^{+1} d\cos\theta P_\ell(\cos\theta) \langle \Delta T(\hat{n})\Delta T(\hat{n}')\rangle, \tag{2.23}$$

where in the last step we took into account that the temperature distribution is assumed to be statistically isotropic.

This two-point function of curvature perturbations originated during inflation (2.20) is related to the correlation of temperature fluctuations of the CMB, in the regime of very low values of ℓ, via the Sachs-Wolfe effect (see, e.g., [189])

$$\langle \Delta T(\vec{n})\Delta T(\vec{n}')\rangle_{SW} = \frac{T_0^2}{25} \langle \mathcal{R}(r_L\vec{n})\mathcal{R}(r_L\vec{n}')\rangle. \tag{2.24}$$

The coefficients C_ℓ are thus predicted to be

$$C_\ell^{SW} = \tfrac{2\pi T_0^2}{25} \int_{-1}^{1} d\cos\theta P_\ell(\cos\theta) \langle \mathcal{R}(r_L\vec{n})\mathcal{R}(r_L\vec{n}')\rangle\ . \tag{2.25}$$

Therefore, the low multipole coefficients, dominated by the Sachs-Wolfe effect, are proportional to the integral

$$C_\ell^{SW} \propto \int_{-1}^{1} dy (1-y)^{\frac{1-n}{2}} P_\ell(y)\ , \tag{2.26}$$

with $y \equiv \cos\theta$. This integral can be computed analytically [8, 105], and we finally find

$$C_\ell^{SW} \;=\; \frac{8\pi T_0^2}{25}\frac{4\pi G}{\epsilon}\frac{H^2(1-\epsilon)^2}{16\pi^2}\frac{\Gamma(3-n)\Gamma(\ell+\frac{n-1}{2})}{\Gamma(\ell+2-\frac{n-1}{2})}\bar{r}_L^{1-n} \,, \tag{2.27}$$

in exact agreement with the result obtained with the momentum-space power spectrum $P_s(k) = |N|^2 k^{n-1}$ [189], by using the formula

$$C_{\ell,SW} = \frac{16\pi^2 T_0^2}{25}\int_0^\infty \frac{dk}{k} P_s(k) j_\ell^2(k r_L) \,, \tag{2.28}$$

and the amplitude $|N|^2$ given by $|N|^2 = \frac{4\pi G}{\epsilon}\frac{H^2(1-\epsilon)^2}{16\pi^2}\frac{2^{3-n}}{\pi^2}\Gamma\left(2-\frac{n}{2}\right)^2\left(\frac{\bar{r}_L}{r_L}\right)^{1-n} \sim \frac{8\pi G H^2}{32\pi^3\epsilon}$. For completeness, taking approximately $n \approx 1$ in (2.27) and using the standard assumption $\bar{r}_L^{(1-n)} \approx O(1)$ [189], the estimated order of magnitude for the amplitude of C_ℓ^{SW} is

$$\ell(\ell+1)C_\ell^{SW} \approx \frac{2GH^2T_0^2}{25\epsilon}\bar{r}_L^{1-n} \sim \frac{2GH^2T_0^2}{25\epsilon} \,. \tag{2.29}$$

We will go back to this point at the end of section 2.4.

We note that if the coefficients C_ℓ^{SW} in (2.25) were actually evaluated using the exact expression (2.16) for the two-point function $\langle\mathcal{R}(t,\vec{x})\mathcal{R}(t,\vec{x}')\rangle$, the integral (2.25) would have been divergent [due to the UV divergences of (2.16) as points $\vec{x}$ and $\vec{x}'$ merge]. This is essentially the feature that we commented in section 2.1. The use of the large distance behavior (2.17 and 2.20) everywhere in the integral (2.25) bypasses the UV divergences and makes the integral convergent. We will see in section 2.4 how the use of a "renormalized" form of the two-point correlator $\langle\mathcal{R}(t,\vec{x})\mathcal{R}(t,\vec{x}')\rangle$ does the same job, but with a slightly different final result for the integral. In a certain limit both results eventually agree, but in general we find a difference that could be potentially probed by observations.

2.3 Spacetime correlators at very short distances

From now on we shall use ϕ to denote both scalar and tensorial fluctuations, and ν to represent the corresponding Hankel index. When necessary we shall return to the letter $\mathcal{R}$ to make some statements explicit.

The two-point function $\langle\phi(t,\vec{x}),\phi(t,\vec{x}')\rangle$ can be expanded at short distances as

$$\langle\phi(t,\vec{x})\phi(t,\vec{x}')\rangle = \frac{H^2(1-\epsilon)^2}{16\pi^2}\left\{\frac{4}{\Delta\vec{x}^2} + \left(\frac{1}{4}-\nu^2\right)\left(-1+2\gamma+\psi(3/2-\nu)+\psi(3/2+\nu)+\log\frac{\Delta\vec{x}^2}{4}\right) + O(\Delta\vec{x}^2)\right\} \tag{2.30}$$

where we have introduced the dimensionless quantity $\Delta\bar{x} \equiv H(1-\epsilon)a\Delta x$. An additional prefactor, $4\pi G/\epsilon$ or $16\pi G$, needs to be included in considering scalar or tensorial perturbations, respectively.

It is clear then that, as advanced in (2.7), the scalar two-point function diverges when the two points merge, $\theta \to 0$. This produces an ultraviolet divergence in the expression for multipole coefficients C_ℓ when evaluated in the Sachs-Wolfe regime, for which one makes the identification (2.24):

$$C_\ell^{SW} = \frac{2\pi T_0^2}{25}\int_{-1}^{1} d\cos\theta P_\ell(\cos\theta)\langle\mathcal{R}(t,\vec{x})\mathcal{R}(t,\vec{x}')\rangle = \infty \ , \tag{2.31}$$

One could cure this divergence by subtracting the appropriate counterterm, as it is normally done in renormalization. Different methods can be used to obtain the subtractions terms. A preferred method for our purposes is the point-splitting version of the adiabatic regularization scheme. The method determines the subtraction terms univocally. The replacement $\langle\mathcal{R}(t,\vec{x})\mathcal{R}(t,\vec{x}')\rangle \to [\langle\mathcal{R}(t,\vec{x})\mathcal{R}(t,\vec{x}')\rangle - G_{Ad}^{(0)}((t,\vec{x}),(t,\vec{x}'))]$, with $G_{Ad}^{(0)}((t,\vec{x}),(t,\vec{x}'))$ being the leading order term in the short-distance expansion (2.30), yields

$$C_\ell^{SW(n=0)} = \frac{2\pi T_0^2}{25}\int_{-1}^{1} d\cos\theta P_\ell(\cos\theta)[\langle\mathcal{R}(t,\vec{x})\mathcal{R}(t,\vec{x}')\rangle - G_{Ad}^{(0)}((t,\vec{x}),(t,\vec{x}'))] < \infty \ . \tag{2.32}$$

This, eventually, produces (2.27) in the large distances limit. It turns out then that, implicitly, the result in the literature is the result of "renormalizing" the two-point function at zero adiabatic order.

Even though this procedure provides a finite result, what we have really encountered in (2.30) is the typical quadratic and logarithmic short-distance behavior of a two-point function in a curved background. So, as argued before, it seems natural to remove both divergences, not only the leading one. Accordingly, it is natural to propose the following identification

$$\langle\Delta T(\vec{n})\Delta T(\vec{n}')\rangle_{SW} = \frac{T_0^2}{25}[\langle\mathcal{R}(t,\vec{x})\mathcal{R}(t,\vec{x}')\rangle - G_{Ad}^{(2)}((t,\vec{x}),(t,\vec{x}'))] \ , \tag{2.33}$$

instead of $\langle\Delta T(\vec{n})\Delta T(\vec{n}')\rangle_{SW} = \frac{T_0^2}{25}[\langle\mathcal{R}(t,\vec{x})\mathcal{R}(t,\vec{x}')\rangle - G_{Ad}^{(0)}((t,\vec{x}),(t,\vec{x}'))]$, as used in (2.32) to obtain (2.27) in the large distances limit. By doing this subtraction, we make the two-point function match the mean-square fluctuation $\langle\mathcal{R}^2(t,\vec{x})\rangle$.

The point-splitting adiabatic method determines the subtraction terms univocally. The second order counterterm in the short-distance asymptotic expansion is

$$G_{Ad}^{(2)}((t,\vec{x}),(t,\vec{x}')) = \frac{H^2(1-\epsilon)^2}{16\pi^2}\left\{\frac{4}{\Delta\bar{x}^2} + \left(\frac{1}{4}-\nu^2\right)\log\frac{\Delta\bar{x}^2}{4} + \frac{2-\epsilon}{3(1-\epsilon)^2} + \left(\frac{1}{4}-\nu^2\right)\left(2\gamma + \log\frac{\mu^2}{H^2(1-\epsilon)^2}\right)\right\} \ ,$$

where μ is a renormalization scale and the corresponding prefactor mentioned above for scalar or tensorial perturbations must be considered. We observe immediately that the UV divergences of the self-correlator cancel exactly and we are left with

$$\langle\phi^2(t,\vec{x})\rangle_{ren} = \frac{H^2(1-\epsilon)^2}{16\pi^2}\left\{\left(\frac{1}{4}-\nu^2\right)\left(-1+\psi(3/2-\nu)+\psi(3/2+\nu)-\log\frac{\mu^2}{H^2(1-\epsilon)^2}\right)-\frac{2-\epsilon}{3(1-\epsilon)^2}\right\}. \tag{2.34}$$

The above self-correlators quantify the amplitude of perturbations at a given spacetime point.

2.4 Modified two-point functions and angular power spectrum

In previous sections we have studied the correlator $\langle\phi(t,\vec{x})\phi(t,\vec{x}')\rangle$ and self-correlator $\langle\phi^2(t,\vec{x})\rangle$ of tensorial and scalar perturbations in slow-roll inflation. For an ordinary quantum mechanical system, with a finite number N of degrees of freedom, expectation values of the form $\langle\phi(i)\phi(j)\rangle$ and $\langle\phi^2(i)\rangle$ match when $j=i$ [for instance, in a chain of spins with $\phi(i)\equiv S_z(i)$ and $i=1,\ldots,N$]. However, we are facing here a field theory (with an infinite number of degrees of freedom), and the above matching is not *a priori* guaranteed. This is so because the self-correlator requires renormalization. We may either assume this discontinuity or modify the two-point correlation function to force it to match $\langle\phi^2(t,\vec{x})\rangle_{ren}$ in the coincidence limit $\vec{x}'\to\vec{x}$ [157]. This second possibility was indirectly explored in [146] by analyzing the power spectrum of perturbations in momentum space. It has been somewhat debated in the literature and properly reviewed in [41]. One could naturally argue that the two-point correlator has a well-defined definition in the distributional sense and there is not a mathematical need for any regularization [41, 89]. However, as the spatial points approach each other, the two-point correlator will grow without bound and diverge as the points merge. Therefore, from the physical point of view it seems reasonable to use a regularized form of the two-point correlator to consistently match the self-correlator at coincidence [157] [1]. In the conventional approach the expectation value of the self-correlator $\langle\phi^2(t,\vec{x})\rangle$ plays almost no role. We assume here that the (renormalized) self-correlator is actually playing a physical role (as in the Casimir effect). As we will see shortly, the regularized form of the two-point correlator makes the desired integral (2.25) UV convergent. The consequences of this merit to be explored. Therefore,

[1]Alternatively, one could also consider short-distance modifications of the two-point function due to quantum gravity effects. See, for instance, [144]

we further analyze here this possibility taking advantage of the spacetime viewpoint sketched above.

We shall modify the correlators by adding the subtraction terms prescribed by renormalization and according to (2.33). We note that a distinguishing characteristic of adiabatic renormalization is that the subtraction terms $G^{(2)}_{Ad}((t,\vec{x}),(t,\vec{x}'))$ are well-defined for arbitrary point separation. In general this is not possible for an arbitrary spacetime, but for the homogeneous spaces relevant in cosmology the adiabatic subtraction terms extend to arbitrary large distances. With this in mind, we will finally compute the angular power spectrum for primordial perturbations using the modified spacetime correlators.

As a previous step we will compute the two-point function at leading order in slow-roll.

Two-point function at leading order in slow-roll

The procedure is similar for scalar and tensorial fluctuations, so we will do a general treatment. First, we start off splitting Eqs. (2.12) and (2.16) as a combination of two hypergeometric functions. To this end we use the transformation properties of hypergeometric functions [8]

$$\begin{aligned} F\left(\frac{3}{2}+\nu,\frac{3}{2}-\nu,2,1-Z\right) &= \frac{Z^{-\frac{3}{2}-\nu}\Gamma(-2\nu)}{\Gamma(\frac{3}{2}-\nu)\Gamma(\frac{1}{2}-\nu)} Re\left\{F\left(\frac{3}{2}+\nu,\frac{1}{2}+\nu,1+2\nu,\frac{1}{Z}\right)\right\} \\ &+ \frac{Z^{-\frac{3}{2}+\nu}\Gamma(2\nu)}{\Gamma(\frac{3}{2}+\nu)\Gamma(\frac{1}{2}+\nu)} Re\left\{F\left(\frac{3}{2}-\nu,\frac{1}{2}-\nu,1-2\nu,\frac{1}{Z}\right)\right\} \end{aligned} \tag{2.35}$$

with $Z=\Delta\bar{x}^2/4 \geq 0$. We now expand expression (2.35) as a power series of the "slow-roll" parameter ν around $\nu = 3/2$, and stay at first order (for details see the Appendix). Grouping terms, we arrive at the following expression for the two-point function:

$$\langle\phi(x)\phi(x')\rangle \approx \frac{H^2(1-\epsilon)^2}{16\pi^2}\left\{\frac{4}{\Delta\bar{x}^2} - 2\log\frac{\Delta\bar{x}^2}{4} - 1 + \frac{2}{(3/2-\nu)}\left(\frac{\Delta\bar{x}^2}{4}\right)^{\nu-3/2} + 4Re\left[\log\left(\frac{\Delta\bar{x}}{2}+\sqrt{\frac{\Delta\bar{x}^2}{4}-1}\right)\right]\right\} \tag{2.36}$$

Notice that the UV divergences are just the same as those found in (2.30), but now they are obtained at leading order in the slow-roll expansion. We recover exactly expression (2.30) taking the limit $\Delta\bar{x} \to 0$ and the slow-roll approximation.

Modified two-point function

We can now proceed to do the subtraction. The modified two-point function then reads

$$\langle\phi(x)\phi(x')\rangle_{ren} \approx \frac{H^2(1-\epsilon)^2}{16\pi^2}\left\{\frac{2}{(3/2-\nu)}\left(\frac{\Delta\bar{x}^2}{4}\right)^{\nu-3/2} + 4Re\left[\log\left(\frac{\Delta\bar{x}}{2}+\sqrt{\frac{\Delta\bar{x}^2}{4}-1}\right)\right] - \frac{5}{3} + 4\gamma + 2\log\frac{\mu^2}{H^2}\right\} \tag{2.37}$$

We remark that, at leading order in the slow-roll expansion, this is an expression valid for small and large separations. For scales larger than the Hubble horizon, $\Delta\bar{x} \gg 1$, we can further take the approximation, $4Re\left[\log\left(\frac{\Delta\bar{x}}{2}+\sqrt{\frac{\Delta\bar{x}^2}{4}-1}\right)\right] \approx 2\log(\Delta\bar{x}^2)$.

Physical angular power spectrum

We now compute the corresponding angular power spectrum from the modified two-point function for scalar perturbations and for low multipoles

$$C_\ell^{SW} = \frac{2\pi T_0^2}{25}\int_{-1}^{1}\langle\mathscr{R}(x)\mathscr{R}(x')\rangle_{ren}(y)P_\ell(y)dy \; . \qquad (2.38)$$

By construction this is a finite quantity, even without taking the large separation limit for the two-point function [as it was assumed in going from (2.25) to (2.26) and (2.27)]. To evaluate the logarithmic contributions of (2.37) to (2.38) we take into account that $\int_{-1}^{1} dy \log(1-y)P_\ell(y) = -2/\ell(\ell+1), \ell = 1,2,\dots$. The final result for the angular power spectrum with the modified two-point function is very well approximated by the following analytical expression:

$$C_\ell^{SW} \approx \frac{4\pi G}{\epsilon}\frac{8\pi T_0^2}{25}\frac{H^2(1-\epsilon)^2\bar{r}_L^{1-n}}{16\pi^2}\left\{\frac{\Gamma(\ell+\frac{n-1}{2})}{\Gamma(\ell+2-\frac{n-1}{2})} - \frac{\bar{r}_L^{n-1}}{\ell(\ell+1)}\right\} , \qquad (2.39)$$

where we have used $\nu - \frac{3}{2} = \frac{1-n}{2}$, and n represents the scalar index of inflation $n = 1-4\epsilon-2\delta+O(\epsilon,\delta)^2$. Also notice that expression (2.39) is valid for $\ell \geq 1$, as for $\ell = 0$ there would be present all the constant contributions from the renormalized two-point function (2.37), including the one depending on the renormalization scale. In fact, the renormalization scale may be fixed by imposing the natural condition $C_0^{SW} = 0$.

Notice that the first term in (2.39) reproduces the standard result (2.27), and would have been the result of renormalizing at zeroth adiabatic order. The second one comes from the subtraction terms that we have added to the two-point correlator to continuously match the self-correlator at coincidence, but it shows scale invariance as well. Therefore, Eq. (2.39) is consistent with observations [166].

However, the two terms in (2.39) are competing, and the resulting amplitude for the coefficients C_ℓ^{SW} depends on the instant of time one evaluates $\bar{r}_L$. The first term is proportional to $H^2(t)\bar{r}_L^{(1-n)}(t)/\epsilon(t)$, and it is time independent. However, the second term depends slightly on time. The value of $\bar{r}_L$ varies along the inflationary period, ranging from $\bar{r}_L \approx 1$, immediately after the instant of time t_i at which the scale r_L crosses the Hubble horizon $[a(t_i)r_L \approx H^{-1}(t_i)]$, to $\bar{r}_L \approx e^{60}$, at the end of inflation (we have assumed that inflation lasts for around $N = 60$ e-foldings since the scale r_L exited the horizon at t_i). In the former case, $\bar{r}_L \approx 1$, the amplitude is severely reduced. In the latter situation, $\bar{r}_L^{(n-1)} \sim 10^{-1}$, where we have assumed that $n \approx 0.96$ [166], and the amplitude is then reduced at least 10%. The adequate value of $\bar{r}_L$ to properly evaluate the resulting amplitude in (2.39) is unclear. This question is closely related to the so-called “quantum-to-classical transition" [128], characterizing the period of time at which the primordial quantum perturbations behave as classical ones and define the initial conditions for the postinflationary evolution, along with its associated power spectrum. In momentum space (mode-by-mode picture) this process is thought to happen a few Hubble times after horizon exit [128], when the modes are frozen as classical perturbations. It seems natural to evaluate $\bar{r}_L$ during this period, where quantum fluctuations are imprinted as classical perturbations. However, this quantum-to-classical mechanism is poorly understood, and it has not been rigorously established in the literature. Therefore one may regard $\bar{r}_L^{(n-1)} \equiv \alpha$ as a phenomenological parameter, varying in the range $1 > \alpha > 0$. Note that in the limiting case $\alpha \to 0$ one recovers the standard prediction, and this happens when the subtraction term is evaluated after inflation. If the subtraction terms are evaluated a few e-foldings after the horizon exit of the scale r_L, the parameter α approaches 1 and the physical significance of the correction increases.

This parameter has influence on the relative strength between multipole amplitudes

$$\frac{\ell_2(\ell_2+1)C_{\ell_2}^{SW}}{\ell_1(\ell_1+1)C_{\ell_1}^{SW}} = \frac{\ell_2(\ell_2+1)\frac{\Gamma(\ell_2+\frac{n-1}{2})}{\Gamma(\ell_2+2-\frac{n-1}{2})} - \alpha}{\ell_1(\ell_1+1)\frac{\Gamma(\ell_1+\frac{n-1}{2})}{\Gamma(\ell_1+2-\frac{n-1}{2})} - \alpha} . \tag{2.40}$$

Observations may properly fit the value of this parameter. It produces an observable effect for a significant range of values of α. As remarked above, the details of how this “quantum-to-classical transition” takes place are not well established in the literature, and further work is needed to fully understand this process. Within the present understanding of quantum gravity it is difficult to determine theoretically the value of α and hence the relative impact of the subtraction term in the observed angular power spectrum. However, as we showed above it can potentially be tested with observations.

2.5 Final comments

We have analyzed two-point correlators and self-correlators of primordial cosmological perturbations in quasi-de Sitter spacetime backgrounds, and studied in considerable detail the evaluation of the multipole coefficients C_ℓ of the angular power spectrum. For large separations two-point correlators exhibit nearly scale invariance in a very elegant way. We have deformed the two-point correlators to smoothly match the self-correlators at coincidence with the aim of obtaining a finite quantity for C_ℓ. To this end we have used renormalization methods in homogeneous backgrounds. We stressed the fact that the standard result in the literature is equivalent to a renormalization of the two-point function at zero adiabatic order, and argued that renormalization at second adiabatic order would be more suitable from a physical point of view.

We have studied the physical consequences for the angular power spectrum at low multipoles, i.e. in the so-called Sachs-Wolfe regime of the spectrum. This may change significantly the predictions of inflation, provided the renormalization subtraction terms are evaluated a few e-foldings after the first horizon crossing of the scale r_L. If one accepts a mismatch between the standard two-point correlators and the self-correlators and keeps only the large-scale behavior, the conventional predictions remain unaltered.

We finally stress the importance of getting a better understanding of how to renormalize cosmological observables. The analysis carried out in the spacetime framework for the tree-level power spectrum may offer a way to experimentally probe this issue.

2.6 Appendix

In this short appendix we give the basic steps to obtain the result (2.36). We consider (2.35) first. Since the first prefactor is of order $O((\frac{3}{2}-\nu)^1)$, we only need the corresponding hypergeometric function to be of order $O((\frac{3}{2}-\nu)^0)$. One can see that

$$Z^{-\frac{3}{2}-\nu} Re\left\{F\left(\frac{3}{2}+\nu,\frac{1}{2}+\nu,1+2\nu,\frac{1}{Z}\right)\right\}\Big|_{\nu=3/2} = 6Re\{\log(Z-1)\} - 6\log(Z) + \frac{3}{Z} \quad - \quad \frac{3}{(1-Z)}$$

On the other hand, the second prefactor of (2.35) is of order $O((\frac{3}{2}-\nu)^0)$, so it is necessary to evaluate the second hypergeometric function at first order in the slow-roll series. To this end we will employ the following relation [8]:

$$F\left(\frac{3}{2}-\nu,\frac{1}{2}-\nu,1-2\nu,\frac{1}{Z}\right) = \left(1-\frac{1}{Z}\right)^{-3/4} P^{\nu}_{1/2}\left[\frac{2Z-1}{2\sqrt{Z(Z-1)}}\right] \quad 2^{-2\nu} \quad \Gamma(1-\nu) Z^{-\nu} \ ,$$

together with

$$P^{\nu}_{1/2}(Z) = \left(\frac{Z+1}{Z-1}\right)^{\nu/2} \frac{F\left(-\frac{1}{2},\frac{3}{2},1-\nu,\frac{1-Z}{2}\right)}{\Gamma(1-\nu)} \ . \tag{2.41}$$

At this point one can expand

$$F\left(-\frac{1}{2},\frac{3}{2},1-\nu,\frac{1-Z}{2}\right) \approx F\left(-\frac{1}{2},\frac{3}{2},-\frac{1}{2},\frac{1-Z}{2}\right) + \left(\nu-\frac{3}{2}\right)\frac{dF\left(-\frac{1}{2},\frac{3}{2},1-\nu,\frac{1-Z}{2}\right)}{d\nu}\Big|_{\nu=3/2} , \tag{2.42}$$

where the derivative can be performed using the representation series of the hypergeometric function. Doing all the calculation properly one finally arrives at the following result:

$$
\begin{aligned}
Re\left\{F\left(\frac{3}{2}-\nu,\frac{1}{2}-\nu,1-2\nu,\frac{1}{Z}\right)\right\} \quad &\approx \quad 1 \qquad (2.43)\\
+\left(\frac{3}{2}-\nu\right)\left[\frac{1}{4Z}+\frac{1}{4(1-Z)}-\frac{1}{2}Re\{\log(Z-1)\} \quad &- \quad \frac{1}{2}\log(Z)+2Re\left\{\log\left(\sqrt{Z}+\sqrt{Z-1}\right)\right\}\right] \; .
\end{aligned}
$$

Taking all these results together for $Z\equiv\Delta\bar{x}^2/4$ in (2.35) we can approximate the two-point function as in (2.36). We have also checked numerically that this expansion works well irrespectively of the value of Z.

CHAPTER 3

Renormalized stress-energy tensor for spin-1/2 fields in expanding universes

We provide here an explicit expression for the renormalized vacuum expectation value of the stress-energy tensor $\langle T_{\mu\nu}\rangle$ of a spin-1/2 field in a spatially flat FLRW universe. Its computation is based on the extension of the adiabatic regularization method to fermion fields introduced recently in the literature. We analyze first the case in which the field is free, and then extend the method to include the Yukawa interaction between quantized Dirac fermions and a homogeneous background scalar field. The resulting tensor is given in terms of UV-finite integrals in momentum space, which involve the mode functions that define the quantum state.

As illustrative examples of the method efficiency for the free field case, we see how to compute the renormalized energy density and pressure in two cosmological scenarios of physical interest: a de Sitter spacetime and a radiation-dominated universe. In the second case, we explicitly show that the late-time renormalized stress-energy tensor behaves as that of classical cold matter.

When the Yukawa interaction is switched on, explicit expressions for both $\langle T_{\mu\nu}\rangle$ and the bilinear $\langle\bar{\psi}\psi\rangle$ are derived. These are basic ingredients in the semiclassical Einstein's field equations of fermionic matter in curved spacetime interacting with a background scalar field. The ultraviolet subtracting terms of the adiabatic regularization can be naturally interpreted as coming from appropriate geometric counterterms of the background fields. To test our approach we determine the contribution of the Yukawa interaction to the conformal anomaly in the massless limit and show its consistency with the heat kernel method using the effective action.

In a final stage, we prove that adiabatic regularization and DeWitt-Schwinger point-splitting provide the same result when renormalizing expectation values of the stress-energy tensor for spin-1/2 fields. This generalizes the equivalence found for scalar fields, which is here recovered in a different way. We also argue that the coincidence limit of the DeWitt-Schwinger proper time expansion of the two-point function agrees exactly with the analogous expansion defined by the adiabatic regularization method at any order (for both scalar and spin-1/2 fields). We illustrate the power of the adiabatic method to compute higher order DeWitt coefficients in FLRW universes.

The content of this chapter is fully based on papers [64, 66, 68].

3.1 Context, motivation and specific goals

The quantization of the gravitational interaction is one of the most important and difficult problems in theoretical physics. Quantum field theory in curved spacetime offers a first step to join Einstein's theory of general relativity and quantum field theory in Minkowski space within a self-consistent and successful framework [48, 94, 157, 185].

A major problem in the theory of quantized fields in curved spacetimes is the computation of the expectation values of the stress-energy tensor components. These calculations are rather convoluted, as they involve products of fields at coincident spacetime points, which are ultraviolet (UV) divergent even for free fields. In cosmological scenarios, this is connected to the fundamental phenomenon of particle creation by time-dependent backgrounds [147–150] (see [151] for a historical review).

A very efficient renormalization method, specifically constructed to deal with the UV divergences of a free field in an expanding universe, is adiabatic regularization. Originally, this technique was introduced to tame the divergences of the mean particle number of a scalar field in a FLRW universe [147], and was later extended to get rid of the divergences of the stress-energy tensor in such a way that locality and covariance of the overall renormalization procedure are fully respected [152] [95] [96] [53] [25]. The key ingredient of the adiabatic scheme is the asymptotic expansion of the field modes, in which increasingly higher-order terms in the expansion involve increasingly higher-order time derivatives of the metric (the scale factor). Due to dimensional reasons, this is equivalent to an UV asymptotic expansion in momenta. This way, one can expand adiabatically the integrand of the unrenormalized bilinear, identify the UV-divergent terms, and subtract them directly to obtain a finite, covariant expression. The renormalized expectation value is hence expressed as a finite integral in momentum space, depending exclusively on the mode functions defining the quantum state.

It should be remarked that, in cosmological perturbation theory, other fields apart from the inflaton, such as the Dirac spinor presented here, are regarded themselves as first order (otherwise they would provide inhomogeneities at zeroth order). Then, in order to stay at linear order in cosmological fluctuations at the level of the equations of motion, it is necessary to study these fields as propagating in the perfectly homogeneous FLRW background. This is what it is customary done in fact for the scalar curvature and tensor metric perturbations. Interactions between these metric perturbations and the Dirac field are of higher order in cosmological perturbation theory. Therefore, when adiabatic regularization (or any other renormalization method) is applied to inflationary cosmology, the additional fields should be considered, at leading order, as quantum fields propagating in the homogeneous FLRW spacetime. This is the reason for which the usual metric perturbations shall not be considered along the chapter.

In this chapter we shall be concerned with three main goals, described below.

Adiabatic renormalization for a free fermion field.

One of the main issues with the renormalization program in curved spacetime is that these methods have been mainly developed for free scalar bosons, and less work has been done for other fields. In particular, an adiabatic regularization method for spin one-half fields in an expanding universe was missing until very recently [126, 127]. One of the key features of this extended method is that , while for free scalar fields the well-known WKB expansion[1] provides an adequate solution (see, for instance, [48, 94, 157]), for spin-1/2 fields, however, the adiabatic expansion takes a different form. The method is specially suitable for numerical calculations, [26, 27, 42, 110, 117, 118, 159] and also for analytic approximations [27, 137]. The adiabatic regularization has also been used to scrutinize the two-point function defining the variance and power spectrum in inflationary cosmology and related issues [17–19, 21, 22, 41, 65, 78, 89, 132, 146, 186].

We begin this chapter by applying the adiabatic regularization method to obtain a general and explicit expression for the renormalized stress-energy tensor of a spin

[1]Named for G. Wentzel, H. Kramers, L Brillouin, it is a method for getting approximate solutions to linear differential equations with varying coefficients in 1D problems. This is the case for a time-independent Schrodinger-like equation with a potential function. The idea of the WKB method is that this function can be thought of as being "slowly varying".

one-half field in a FLRW universe. This will be mostly covered by sections 3.2, 3.3, 3.4 and Appendices A, B and C. The main result, given in equations (3.69), (3.72) and (3.73), is written in terms of UV-convergent momentum integrals involving the field modes. This is a necessary and unavoidable step to prepare the method to be used for numerical computations in cosmology. As illustrative examples, we study the renormalized stress-energy tensor in de Sitter space and in a radiation-dominated universe. In both examples, we need to specify appropriate initial conditions in order to ensure the renormalizability properties of the tensor. We also prove here that the same procedure used in [126, 127] to obtain the adiabatic expansion of the fermionic field modes leads to the well-known WKB-type expansion when the algorithm is applied for scalar modes. This confirms definitely the appropriateness of the fermionic adiabatic regularization method.

Adiabatic renormalization for a fermion field interacting with a background scalar field trough Yukawa coupling

Particle creation also takes place if the quantized field, either a boson or a fermion, is coupled to a classical background scalar field evolving non-adiabatically in time. In this case, the interaction term acts in the boson/fermion equation of motion as a time-dependent effective mass, which excites the field and increases its mean particle number. The most paradigmatic example of this is probably preheating after inflation [123, 124]. In the same way as before, new UV divergences associated to this particle creation appear in the expectation values of the different bilinears, not present in the absence of interaction, which must be appropriately removed to obtain finite quantities. Although renormalization of expectation values for interacting fields is generally much more complicated, adiabatic regularization can be generalized to include interactions to classical scalar background fields. In this case, the adiabatic expansion of the field modes used to identify the UV- divergent terms depend on both the scale factor and the background field, as well as their respective time-derivatives. If the quantized field is a scalar with a Yukawa-type coupling, the adiabatic expansion is still of the WKB form [137][28]. However, a generalization of the adiabatic scheme for other interacting species in an expanding universe is absent in the literature. Here, we will try to partially fill in this gap.

In this chapter we shall also extend the adiabatic regularization method to Dirac fields living in a FLRW universe and interacting, via the standard Yukawa coupling, with an external scalar field. This will be covered in sections 3.5, and appendices D, E and F. In this approach, the Dirac field is quantized, while both the metric and the background field are regarded as classical. This kind of system appears for example in fermionic preheating, in which the inflaton acts as a background scalar field oscillating around the minimum of its potential, and decays non-perturbatively into fermions due to its Yukawa interactions [36, 98, 104, 106, 107, 160]. Another example is the decay of the Standard Model (SM) Higgs after inflation, in which the Higgs condensate oscillates around the minimum of its potential, and transfers part of its energy into all the massive fermions of the Standard Model, coupled to the Higgs with the usual Standard Model Yukawa couplings [86] [82] (another part being transferred to the SM gauge bosons [82] [83, 88]). In certain models, the Higgs decay may also lead to the reheating of the Universe [46, 87, 97]. In this work, we will not focus on a particular scenario, but consider arbitrary time-dependent scale factors and background fields. The main objective is to provide well-motivated and rigorous expressions for the renormalized expectation values of the fermion stress-energy tensor $\langle T_{\mu\nu}\rangle$ and the bilinear $\langle\bar{\psi}\psi\rangle$. In the semiclassical equations of motion, these are the quantities that incorporate the backreaction of the created matter onto the background fields. To check the validity of the adiabatic method, we will also compute the contribution of the Yukawa interaction to the conformal anomaly in the massless limit, and check its consistency with the heat kernel method using the effective action.

Equivalence between adiabatic and DeWitt-Schwinger

An alternative asymptotic expansion (for the two-point function) to consistently identify the subtraction terms in a generic spacetime was suggested by DeWitt [71], generalizing the Schwinger proper-time formalism. The DeWitt-Schwinger expansion was implemented with the point-splitting renormalization technique in [59] and it was nicely rederived from the local momentum-space representation introduced by Bunch and Parker [54]. Furthermore, by brute force calculation Birrell [47] (see also the appendix in [25]) checked that point-splitting and adiabatic renormalization give the same renormalized stress-energy tensor when applied to scalar fields in homogeneous universes.

The extension of the adiabatic regularization method to spin-1/2 fields has been achieved very recently [66, 68, 126, 127] (see also [101, 102]). The main difficulty in extending the adiabatic scheme to fermion fields is that the proper asymptotic adiabatic expansion of the spin-1/2 field modes does not fit the WKB-type expansion, as happens for scalar fields. However, as shown in [66, 68, 126, 127], the method has passed a very nontrivial test of consistency. A major goal of this section is to prove that adiabatic regularization and DeWitt-Schwinger point-splitting will give the same result for the renormalized expectation values of the stress-energy tensor of spin-1/2 fields. We base our proof on the well-known fact that two different methods to compute $\langle T_{\mu\nu}\rangle$ can differ at most by a linear combination of conserved local curvature tensors. This result assumes that the renormalization methods obey locality and covariance [185]. Since $\langle T_{\mu\nu}\rangle$ has dimensions of (length)$^{-4}$ the only candidates are $m^4 g_{\mu\nu}$, $m^2 G_{\mu\nu}$, ${}^{(1)}H_{\mu\nu}$ and ${}^{(2)}H_{\mu\nu}$ (the last two terms can be obtained by functionally differentiating the quadratic curvature Lagrangians R^2 and $R_{\mu\nu}R^{\mu\nu}$). It can be seen that the stress-energy tensor only needs subtraction up to fourth order in the derivatives of the metric [48, 157], so that higher order contributions need not be considered.

Therefore, the possible difference between the expectation values $\langle T_{\mu\nu}\rangle^{Ad}$, computed with adiabatic regularization, and $\langle T_{\mu\nu}\rangle^{DS}$, computed with the (DeWitt-Schwinger) point-splitting method, is parametrized by four dimensionless constants c_i, $i = 1, ...4$.

$$\langle T_{\mu\nu}\rangle^{Ad} - \langle T_{\mu\nu}\rangle^{DS} = c_1 {}^{(1)}H_{\mu\nu} + c_2 {}^{(2)}H_{\mu\nu} + c_3 m^2 G_{\mu\nu} + c_4 m^4 g_{\mu\nu} \ . \tag{3.1}$$

In our case, the constant c_4 is necessarily zero since both prescriptions lead to a vanishing renormalized stress-energy tensor when restricted to Minkowski spacetime. Moreover, in a FLRW space-time the conserved tensors ${}^{(2)}H_{\mu\nu}$ and ${}^{(1)}H_{\mu\nu}$ are not independent, so we can assume without loss of generality that $c_2 \equiv 0$. Therefore, we are left with

$$\langle T_{\mu\nu}\rangle^{Ad} - \langle T_{\mu\nu}\rangle^{DS} = c_1 {}^{(1)}H_{\mu\nu} + c_3 m^2 G_{\mu\nu} \ . \tag{3.2}$$

Moreover, taking traces in the above relation we get

$$\langle T\rangle^{Ad} - \langle T\rangle^{DS} = -6c_1 \Box R - c_3 m^2 R \ . \tag{3.3}$$

In the massless limit, the classical action of the spin-1/2 field is conformally invariant. The trace anomaly calculated with the new adiabatic regularization method has been proved to be in exact agreement with that obtained by other renormalization methods, and in particular with the DeWitt-Schwinger point-splitting method. This implies that $c_1 = 0$. Obviously, the same arguments and conclusions apply for a scalar field. The equivalence between both methods is therefore reduced to check that the remaining parameter c_3 is also zero. This is actually the most subtle point.

The comparison between $\langle T^{Ad}\rangle$ and $\langle T^{DS}\rangle$ can be better studied by taking into account that, for spin-1/2 fields, $\langle T\rangle = m\langle\bar{\psi}\psi\rangle$. The equivalence is then reduced to prove that

$${}^{(4)}\langle\bar{\psi}\psi\rangle^{Ad} = {}^{(4)}\langle\bar{\psi}\psi\rangle^{DS} \ , \tag{3.4}$$

where ${}^{(4)}\langle\bar{\psi}\psi\rangle^{Ad,DS}$ stands for the subtraction terms, up to fourth order in the derivatives of the metric, in the adiabatic and DeWitt-Schwinger expansions respectively. As remarked above, the fourth order is the order required to remove, in general, the UV divergences in the stress-energy tensor. To prove (3.4) and achieve our goal we will make use of the (Bunch-Parker) local momentum-space representation [54] of the two-point function. A conceptual advantage of our strategy in comparing both renormalization methods is that it offers a better way to spell out their equivalence. In fact, we will also show that the equivalence found at fourth order can be extended to higher order, for both scalar and spin-1/2 fields.

Notation

We use natural units $c = \hbar = 1$, a 4 dimensional spacetime, and the conventions in [48, 157]. Namely, we follow the $(-,-,-)$ convention of [136], in which the metric signature is $(+,-,-,-)$, the Riemann tensor is defined through $[\nabla_\mu, \nabla_\nu]k^\alpha = -R_{\mu\nu\beta}{}^{\alpha}k^\beta$, the Ricci tensor and scalar defined as $R_{\mu\nu} = R^{\alpha}{}_{\mu\alpha\nu}$ and $R = g^{\mu\nu}R_{\mu\nu}$. We work in a spatially flat FLRW background, $ds^2 = dt^2 - a^2(t)d\vec{x}^2$, and we use the Dirac-Pauli representation for the Dirac gamma matrices, $\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$, $\vec{\gamma} = \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix}$, with $\vec{\sigma}$ the usual Pauli matrices.

3.2 Free quantized spin-1/2 fields and the adiabatic expansion

A spin-1/2 field ψ of mass m in curved spacetime is described by the Dirac equation

$$(i\underline{\gamma}^\mu \nabla_\mu - m)\psi = 0 , \tag{3.5}$$

where $\underline{\gamma}^\mu(x)$ are the spacetime-dependent Dirac matrices satisfying the anticommutation relations $\{\underline{\gamma}^\mu, \underline{\gamma}^\nu\} = 2g^{\mu\nu}$, and ∇_μ is the Levi-Civita connection. In a suitable coordinate chart we can write $\nabla_\mu = \partial_\mu - \Gamma_\mu$ where Γ_μ is the 1-form spin-connection.

In a spatially flat FLRW universe, $ds^2 = dt^2 - a^2(t)d\vec{x}^2$, the matrices $\underline{\gamma}^\mu(t)$ are related with the constant Minkowskian matrices γ^α (which satisfy the Clifford algebra $\{\gamma^\alpha, \gamma^\beta\} = 2\eta^{\alpha\beta}$) by $\underline{\gamma}^0(t) = \gamma^0$ and $\underline{\gamma}^i(t) = \gamma^i/a(t)$. On the other hand, the spin connection reads $\Gamma_0 = 0$ and $\Gamma_i = \frac{\dot{a}}{2}\gamma_0\gamma_i$ in this metric. Therefore, $\underline{\gamma}^\mu\Gamma_\mu = -\frac{3\dot{a}}{2a}\gamma_0$, and the differential equation (3.5) can be written as

$$\left(i\gamma^0\partial_0 + \frac{3i}{2}\frac{\dot{a}}{a}\gamma^0 + \frac{i}{a}\vec{\gamma}\cdot\vec{\nabla} - m\right)\psi = 0 , \tag{3.6}$$

where $\vec{\gamma} = (\gamma^1, \gamma^2, \gamma^3)$. Throughout this paper we shall work with the Dirac-Pauli representation for the Dirac matrices

$$\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} , \qquad \vec{\gamma} = \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix} , \tag{3.7}$$

where $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ are the usual Pauli matrices. Since the spacetime background is homogeneous we can work in Fourier space. By extending the quantization procedure in Minkowski space (see for instance [165]) one can construct, for a given $\vec{k}$, two independent spinor solutions as

$$u_{\vec{k}\lambda}(x) = u_{\vec{k}\lambda}(t)e^{i\vec{k}\cdot\vec{x}} = \frac{e^{i\vec{k}\cdot\vec{x}}}{\sqrt{(2\pi)^3 a^3}} \begin{pmatrix} h^I_k(t)\xi_\lambda(\vec{k}) \\ h^{II}_k(t)\frac{\vec{\sigma}\cdot\vec{k}}{k}\xi_\lambda(\vec{k}) \end{pmatrix}, \tag{3.8}$$

where $k \equiv |\vec{k}|$ and ξ_λ is a constant and normalized two-component spinor $\xi^\dagger_\lambda \xi_{\lambda'} = \delta_{\lambda'\lambda}$. In this decomposition, h^I_k and h^{II}_k are two particular time-dependent functions obeying from (3.6) the following coupled differential equations,

$$h^{II}_k = \frac{ia}{k}(\partial_t + im)h^I_k , \qquad h^I_k = \frac{ia}{k}(\partial_t - im)h^{II}_k , \tag{3.9}$$

and the following two uncoupled second order differential equations,

$$\left(\partial_t^2 + \frac{\dot{a}}{a}\partial_t + im\frac{\dot{a}}{a} + m^2 + \frac{k^2}{a^2}\right)h^I_k = 0 , \tag{3.10}$$

$$\left(\partial_t^2 + \frac{\dot{a}}{a}\partial_t - im\frac{\dot{a}}{a} + m^2 + \frac{k^2}{a^2}\right)h^{II}_k = 0 . \tag{3.11}$$

For our purposes, it is convenient to use helicity eigenstates $\xi_\lambda(\vec{k})$, which follow the property $\frac{\vec{\sigma}\vec{k}}{2k}\xi_\lambda(\vec{k}) = \frac{\lambda}{2}\xi_\lambda(\vec{k})$, where $\lambda/2 = \pm 1/2$ represent the eigenvalues for the helicity. Their explicit form is ($\vec{k} = (k_1, k_2, k_3)$)

$$\xi_{+1}(\vec{k}) = \frac{1}{\sqrt{2k(k+k_3)}} \begin{pmatrix} k + k_3 \\ k_1 + ik_2 \end{pmatrix} , \qquad \xi_{-1}(\vec{k}) = \frac{1}{\sqrt{2k(k+k_3)}} \begin{pmatrix} -k_1 + ik_2 \\ k + k_3 \end{pmatrix} . \tag{3.12}$$

Given a particular solution $(h^I_k(t), h^{II}_k(t))$ to the equations (3.9), one can naturally construct a new solution to the same equations $(-h^{II*}_k(t), h^{I*}_k(t))$. In Minkowski space, this is equivalent to jumping from a positive frequency solution to a negative frequency one. Therefore, one can construct two more independent and orthogonal solutions as

$$v_{\vec{k}\lambda}(x) = \frac{e^{-i\vec{k}\cdot\vec{x}}}{\sqrt{(2\pi)^3 a^3}} \begin{pmatrix} -h^{II*}_k(t)\xi_{-\lambda}(\vec{k}) \\ -h^{I*}_k(t)\frac{\vec{\sigma}\cdot\vec{k}}{k}\xi_{-\lambda}(\vec{k}) \end{pmatrix} . \tag{3.13}$$

Using this approach, the $v_{\vec{k}\lambda}(x)$-modes are obtained by the charge conjugation operation: $v_{\vec{k}\lambda} = u^c_{\vec{k}\lambda} \equiv C\left[\bar{u}_{\vec{k}\lambda}\right]^T = i\gamma^2 u^*_{\vec{k}\lambda}$. Note that $-i\sigma_2\xi^*_\lambda(\vec{k}) = \lambda\xi_{-\lambda}(\vec{k})$. It is easy to check that $u^\dagger_{\vec{k}\lambda} v_{-\vec{k}\lambda'} = 0$ and $(u_{\vec{k}\lambda}, v_{\vec{k}'\lambda'}) = 0$, where the Dirac scalar product is given by

$$(\psi_1,\psi_2) = \int d^3x a^3 \psi_1^\dagger \psi_2 \,. \tag{3.14}$$

The normalization condition for the four-spinors, $(u_{\vec{k}\lambda}, u_{\vec{k}'\lambda'}) = (v_{\vec{k}\lambda}, v_{\vec{k}'\lambda'}) = \delta_{\lambda\lambda'}\delta^{(3)}(\vec{k}-\vec{k}')$ leads to

$$|h^I_k(t)|^2 + |h^{II}_k(t)|^2 \;\; = \;\; 1 \,. \tag{3.15}$$

This condition guaranties the standard anticommutation relations for the creation and annihilation operators $B_{\vec{k}\lambda}$ and $D_{\vec{k}\lambda}$, defined by the expansion of the Dirac field in terms of the spinors introduced above

$$\psi(x) = \int d^3\vec{k} \sum_\lambda \left[B_{\vec{k}\lambda} u_{\vec{k}\lambda}(x) + D^\dagger_{\vec{k}\lambda} v_{\vec{k}\lambda}(x)\right] \,. \tag{3.16}$$

Adiabatic expansion

The adiabatic regularization method for spin-1/2 fields, introduced in [126, 127], is based on the following asymptotic UV ansatz for the field modes

$$h^I_k(t) \sim \sqrt{\frac{\omega+m}{2\omega}}\, e^{-i\int^{t'}\Omega(t')dt'} F(t) \quad , \quad h^{II}_k(t) \sim \sqrt{\frac{\omega-m}{2\omega}}\, e^{-i\int^{t'}\Omega(t')dt'} G(t) \,, \tag{3.17}$$

where $\omega \equiv \sqrt{(k/a(t))^2 + m^2}$ is the frequency of the mode and the time-dependent functions $\Omega(t)$, $F(t)$ and $G(t)$ are expanded adiabatically as

$$\begin{aligned} \Omega(t) &= \omega + \omega^{(1)} + \omega^{(2)} + \omega^{(3)} + \omega^{(4)} + \ldots \,, \\ F(t) &= 1 + F^{(1)} + F^{(2)} + F^{(3)} + F^{(4)} + \ldots \,, \\ G(t) &= 1 + G^{(1)} + G^{(2)} + G^{(3)} + G^{(4)} + \ldots \,. \end{aligned} \tag{3.18}$$

Here, $\omega^{(n)}$, $F^{(n)}$ and $G^{(n)}$ are functions of adiabatic order n, which means that they contain n derivatives of the scale factor (for example, $\dot{a}$ is of adiabatic order 1 and $\ddot{a}\dot{a}^2$ is of adiabatic order 4). In the expansions above, we impose $F^{(0)} = G^{(0)} \equiv 1$ and $\omega^{(0)} \equiv \omega$ to recover the Minkowskian solutions in the adiabatic regime.

In order to obtain the different terms of (3.18), we substitute (3.17) into the Dirac equations (3.9) and the normalization condition (3.15). We have then the following system of three equations

$$\begin{aligned}
\Omega F + i\dot{F} + i\frac{F}{2}\frac{d\omega}{dt}\left[\frac{1}{\omega+m} - \frac{1}{\omega}\right] - mF &= (\omega - m)G \ , \\
\Omega G + i\dot{G} + i\frac{G}{2}\frac{d\omega}{dt}\left[\frac{1}{\omega-m} - \frac{1}{\omega}\right] + mG &= (\omega + m)F \ , \\
(\omega+m)FF^* + (\omega-m)GG^* &= 2\omega \ .
\end{aligned} \tag{3.19}$$

We can obtain expressions for $F^{(n)}$, $G^{(n)}$ and $\omega^{(n)}$ by substituting (3.18) into (3.19) and solving the system order by order. In this process, we need to treat independently the real and imaginary parts of $F^{(n)}$ and $G^{(n)}$. The expressions obtained contain ambiguities, which eventually do not appear in the final renormalized physical quantities $\langle\bar{\psi}\psi\rangle$ and $\langle T_{\mu\nu}\rangle$. For sake of simplicity, one can impose order by order the additional simplifying condition $F^{(n)}(m) = G^{(n)}(-m)$, which removes the spurious ambiguities and is a natural choice due to the symmetries of the equations of motion (3.9) under the change of the mass sign. With this, one obtains for the first two orders

$$\omega^{(1)} = 0 \ , \tag{3.20}$$

$$F^{(1)} = -i\frac{m\dot{a}}{4\omega^2 a} \ , \tag{3.21}$$

and

$$\omega^{(2)} = \frac{5m^4\dot{a}^2}{8a^2\omega^5} - \frac{3m^2\dot{a}^2}{8a^2\omega^3} - \frac{m^2\ddot{a}}{4a\omega^3} \ , \tag{3.22}$$

$$F^{(2)} = -\frac{5m^4\dot{a}^2}{16a^2\omega^6} + \frac{5m^3\dot{a}^2}{16a^2\omega^5} + \frac{3m^2\dot{a}^2}{32a^2\omega^4} - \frac{m\dot{a}^2}{8a^2\omega^3} + \frac{m^2\ddot{a}}{8a\omega^4} - \frac{m\ddot{a}}{8a\omega^3} \ . \tag{3.23}$$

The third and fourth order contributions are written in appendix A for completeness.

The adiabatic renormalization method consists in expanding adiabatically the momentum integral of the quantity we want to renormalize using (3.17), and subtracting enough adiabatic terms in order to ensure its convergence in the UV regime. The renormalization of the two-point function at coincidence requires subtraction up to second-order, while the stress-energy tensor needs subtraction up to fourth-order. We apply this procedure to the renormalization of the stress-energy tensor in section 3.3. First, to gain physical intuition, and for readers more familiarized with the WKB-expansion for scalar modes, we see below that a similar technique to the one used here can be equally applied for scalar fields. We rediscover this way the standard WKB-type adiabatic expansion.

Another view on the adiabatic expansion for scalar fields

In this section, we provide a new view on the adiabatic expansion for scalar modes. Mimicking the procedure designed to deal with fermions, we will recover the well-know bosonic WKB adiabatic expansion without assuming it as an a priori input. Since the scalar field is more easy to manage, it can serve to illustrate the prescription used for fermions. A free scalar field of mass m in curved spacetime is described by the wave equation

$$(\nabla^\mu \nabla_\mu + m^2 + \xi R)\phi = 0 \ , \tag{3.24}$$

where ξ is the coupling of the field to the scalar curvature R. Taking the spatially flat FLRW metric and associated coordinate chart, the equation takes the form

$$\frac{\partial^2 \phi}{\partial t^2} + 3\frac{\dot{a}}{a}\frac{\partial \phi}{\partial t} + \frac{1}{a^2}\sum_i \frac{\partial^2 \phi}{\partial {x^i}^2} + (m^2 + \xi R)\phi = 0 \ , \tag{3.25}$$

with $R = 6\frac{\ddot{a}}{a} + 6\frac{\dot{a}^2}{a^2}$. One now expands the field as

$$\phi(\vec{x}, t) = \int \frac{d^3\vec{k}}{\sqrt{2(2\pi a(t))^3}} \left[A_{\vec{k}} e^{i\vec{k}\vec{x}} h_k(t) + A^\dagger_{\vec{k}} e^{-i\vec{k}\vec{x}} h^*_k(t) \right] \ , \tag{3.26}$$

where $h_k(t)$ are time-dependent functions, and the commutation relations for the creation and destruction operators are $[A_{\vec{k}}, A^\dagger_{\vec{k}'}] = \delta^{(3)}(\vec{k} - \vec{k}')$, $[A^\dagger_{\vec{k}}, A^\dagger_{\vec{k}'}] = 0$, and $[A_{\vec{k}}, A_{\vec{k}'}] = 0$. If we substitute (3.26) into (3.25), we find that $h_k(t)$ obeys the differential equation

$$\frac{d^2 h_k}{dt^2} + (\omega_k^2(t) + \sigma)h_k = 0 \ , \tag{3.27}$$

with $\omega_k(t) \equiv \sqrt{(k/a(t))^2 + m^2}$ and $\sigma \equiv (6\xi - 3/4)\frac{\dot{a}^2}{a^2} + (6\xi - 3/2)\frac{\ddot{a}}{a}$. On the other hand, the right normalization condition for scalar fields is

$$h_k \dot{h}^*_k - h^*_k \dot{h}_k = 2i \ . \tag{3.28}$$

As for our analysis for spin-1/2 fields, we assume the following generic ansatz for the adiabatic expansion of the mode functions $h_k(t)$

$$h_k(t) \sim H(t) e^{-i\int^t \Omega(t')dt'} \ , \tag{3.29}$$

where here $H(t)$ and $\Omega(t)$ are real functions. This simplifying assumption is somewhat equivalent to the natural symmetry relation used for spin-1/2 fields $F^{(n)}(m) = G^{(n)}(-m)$. We can expand them adiabatically as

$$H(t) = \frac{1}{\sqrt{\omega_k}} + H^{(1)}(t) + H^{(2)}(t) + H^{(3)}(t) + H^{(4)}(t) + \dots \ , \tag{3.30}$$

and

$$\Omega(t) = \omega_k + \omega^{(1)}(t) + \omega^{(2)}(t) + \omega^{(3)}(t) + \omega^{(4)}(t) + \dots \,. \tag{3.31}$$

As for fermions, we have ensured that the zeroth order term of the expansion recovers the Minkowskian solutions: $H^{(0)}(t) \equiv \omega_k^{-1/2}(t)$ and $\omega^{(0)}(t) \equiv \omega_k(t) \equiv \omega(t)$. By substituting the ansatz (3.29) into the field equation (3.27) and the Wronskian (3.28), we obtain the following system of two equations

$$\begin{aligned} \ddot{H} - H\Omega^2 - 2i\Omega\dot{H} - iH\dot{\Omega} + (\omega^2 + \sigma)H &= 0\,, \\ \Omega H^2 &= 1\,. \end{aligned} \tag{3.32}$$

Substituting (3.30) and (3.31) into (3.32) and solving the system order by order, we find that the first term of both expansions is null, $\omega^{(1)} = H^{(1)} = 0$, and that the second order terms are

$$\omega^{(2)}(t) = \frac{5m^4\dot{a}^2}{8a^2\omega^5} - \frac{m^2\dot{a}^2}{2a^2\omega^3} - \frac{\dot{a}^2}{2a^2\omega} + \frac{3\xi\dot{a}^2}{a^2\omega} - \frac{m^2\ddot{a}}{4a\omega^3} - \frac{\ddot{a}}{2a\omega} + \frac{3\xi\ddot{a}}{a\omega}\,, \tag{3.33}$$

and

$$H^{(2)}(t) = -\frac{5m^4\dot{a}^2}{16a^2\omega^{13/2}} + \frac{m^2\dot{a}^2}{4a^2\omega^{9/2}} + \frac{\dot{a}^2}{4a^2\omega^{5/2}} - \frac{3\xi\dot{a}^2}{2a^2\omega^{5/2}} + \frac{m^2\ddot{a}}{8a\omega^{9/2}} + \frac{\ddot{a}}{4a\omega^{5/2}} - \frac{3\xi\ddot{a}}{2a\omega^{5/2}}\,. \tag{3.34}$$

This algorithm can be extended to all orders. One can immediately check that the expansions obtained this way are equivalent to the usual WKB-type expansions used for scalar fields [152] [95] [96]:

$$h_k(t) \sim \frac{1}{\sqrt{W_k(t)}} e^{-i\int^t W_k(t')dt'}\,. \tag{3.35}$$

More specifically, one confirms that the expansion for $W_k(t) = \omega_k + \omega^{(1)} + \omega^{(2)} + \dots$ is the same as the one for $\Omega_k(t)$ obtained above. One also finds that the $H^{(n)}$ are equal to

$$H^{(n)} = \left(\frac{1}{\sqrt{\omega_k + \omega^{(1)} + \omega^{(2)} + \dots}}\right)^{(n)}\,. \tag{3.36}$$

One rediscovers this way the WKB expansion for scalar fields. The advantage of this strategy has been showed for spin-1/2 fields, as it seems that an efficient WKB-type adiabatic expansion for fermions in an expanding universe does not exist [126, 127].

3.3 Renormalization of the stress-energy tensor

Dirac stress-energy tensor components

The classical stress-energy tensor for a Dirac field in curved spacetime is given by

$$T_{\mu\nu} = \frac{i}{2}\left[\bar{\psi}\underline{\gamma}_{(\mu}\nabla_{\nu)}\psi - \bar{\psi}\overleftarrow{\nabla}_{(\nu}\underline{\gamma}_{\mu)}\psi\right] , \tag{3.37}$$

where ψ is the Dirac field and $\underline{\gamma}(x)$ are the spacetime-dependent Dirac matrices. In the case of a FLRW universe, its homogeneity and spatial isotropy imply that we only have two independent components for this tensor: the energy density, related with the 00-component, and the pressure, related with the ii-component. The 00-component can be written as

$$T_0^0 = \frac{i}{2}\left(\bar{\psi}\gamma^0\frac{\partial\psi}{\partial t} - \frac{\partial\bar{\psi}}{\partial t}\gamma^0\psi\right) , \tag{3.38}$$

while the ii-component is

$$T_i^i = \frac{i}{2a}\left(\bar{\psi}\gamma^i\frac{\partial\psi}{\partial x^i} - \frac{\partial\bar{\psi}}{\partial x^i}\gamma^i\psi\right) , \tag{3.39}$$

(not sum on i implied). The former is directly computed using $\Gamma_0 = 0$ and $\underline{\gamma}_0(x) = \gamma_0$. The latter is obtained taking into account that $\Gamma_i = \frac{\dot{a}}{2}\gamma_0\gamma_i$ and $\underline{\gamma}_i(x) = \gamma_i/a(t)$.

The next step is to compute the formal vacuum expectation values of the quantized stress-energy tensor. To this end, we will use the expansion of the Dirac field (3.16) in terms of the creation and annihilation operators. As a necessary previous result, we first compute the quantity $\langle\bar{\psi}\gamma^\mu\partial_\nu\psi\rangle$. It is given by

$$\langle\bar{\psi}\gamma^\mu\partial_\nu\psi\rangle = \int d^3\vec{k}\sum_{\lambda=\pm1}\left(\bar{v}_{\vec{k}\lambda}\gamma^\mu\partial_\nu v_{\vec{k}\lambda}\right) . \tag{3.40}$$

With this result and with equation (3.13), we will compute the vacuum expectation value of (3.38) and (3.39) and obtain the corresponding expectation values for the energy density and pressure operators.

Renormalized energy density

Let's start with the energy density. If we take the expectation value of (3.38) over the vacuum and use (3.40) and (3.13), we get after some algebra

$$\langle T_{00}\rangle = \frac{1}{2\pi^2a^3}\int_0^\infty dk k^2\rho_k , \tag{3.41}$$

where

$$\rho_k(t) \equiv i\left(h_k^I\frac{\partial h_k^{I*}}{\partial t} + h_k^{II}\frac{\partial h_k^{II*}}{\partial t} - h_k^{I*}\frac{\partial h_k^I}{\partial t} - h_k^{II*}\frac{\partial h_k^{II}}{\partial t}\right) . \tag{3.42}$$

Expression (3.41) contains quartic, quadratic and logarithmic ultraviolet divergences, and consequently, we must expand its integrand adiabatically and subtract from it enough terms of its expansion in order to have a finite quantity. By dimensionality, one expects to need subtraction up to fourth adiabatic order[2]. To see this, we expand (3.42) as

$$\rho_k = \rho_k^{(0)} + \rho_k^{(1)} + \rho_k^{(2)} + \rho_k^{(3)} + \rho_k^{(4)} + \dots \ , \tag{3.44}$$

where $\rho_k^{(n)}$ is of nth adiabatic order. In order to obtain the terms of the expansion, we substitute (3.17) and (3.18) into (3.42) and obtain the contribution from the different adiabatic terms. We find that the zeroth adiabatic term corresponds to the usual Minkowskian divergence

$$\rho_k^{(0)} = -2\omega \ , \tag{3.45}$$

and that the first odd terms are null

$$\rho_k^{(1)} = \rho_k^{(3)} = 0 \ . \tag{3.46}$$

On the other hand, the second order term is given by

$$\rho_k^{(2)} = \frac{\omega+m}{\omega}\left(\Im\mathfrak{m}\dot{F}^{(1)} - |F^{(1)}|^2\omega - 2F^{(2)}\omega - \omega^{(2)}\right) + \frac{\omega-m}{\omega}(F\to G) \ , \tag{3.47}$$

[$(F \to G)$ is the same expression as in the first parenthesis but changing F by G] and the fourth-order term is

$$\begin{aligned}\rho_k^{(4)} = \quad & \frac{\omega+m}{\omega}\Big(\Im\mathfrak{m}\dot{F}^{(3)} - \dot{F}^{(2)}\Im\mathfrak{m}F^{(1)} - (F^{(2)})^2\omega - (F^{(3)*}F^{(1)} + F^{(1)*}F^{(3)} + 2F^{(4)})\omega \\ & + F^{(2)}(\Im\mathfrak{m}\dot{F}^{(1)} - 2\omega^{(2)}) - |F^{(1)}|^2\omega^{(2)} - \omega^{(4)}\Big) + \frac{\omega-m}{\omega}(F\to G) \ .\end{aligned} \tag{3.48}$$

For completeness, the n-th adiabatic order contribution to (3.42) is given by

$$\rho_k^{(n)} = -2\omega(t)^{(n)} + i\frac{\omega+m}{2\omega}\left[F\dot{F}^* - F^*\dot{F}\right]^{(n)} + i\frac{\omega-m}{2\omega}\left[G\dot{G}^* - G^*\dot{G}\right]^{(n)}. \tag{3.49}$$

In order to write (**??**) and (3.48) in terms of ω and the mass, we use (3.20)-(3.23) and (3.296)-(3.299). This gives

$$\rho_k^{(2)} = -\frac{m^4\dot{a}^2}{4\omega^5 a^2} + \frac{m^2\dot{a}^2}{4\omega^3 a^2} \ , \tag{3.50}$$

[2]It follows on dimensional grounds that if a quantity Q has dimensions M^d (where M means mass), the nth adiabatic order term $Q^{(n)}$ in its expansion

$$Q = Q^{(0)} + Q^{(1)} + Q^{(2)} + Q^{(3)} + Q^{(4)} + \dots \ , \tag{3.43}$$

decays in the UV limit as $\mathcal{O}(k^{-\lambda})$ with $\lambda \geq \lambda_* \equiv n-d$. This can be confirmed by just looking at expressions (3.20)-(3.23) and (3.296)-(3.299) . For $Q=\rho_k$, we have dimension $d=1$, and we would require $\rho_k \sim k^{-4}$ in the UV limit in order to have (3.148) finite. Therefore, as $n = \lambda_* + d = 4+1 = 5$, we need to subtract in (3.41) all expansion terms from $\rho_k^{(0)}$ to $\rho_k^{(4)}$. This way, the first contribution to ρ_k comes from the fifth order adiabatic terms.

and

$$
\begin{aligned}
\rho_k^{(4)} &= \frac{105m^8\dot{a}^4}{64w^{11}a^4} - \frac{91m^6\dot{a}^4}{32w^9a^4} + \frac{81m^4\dot{a}^4}{64w^7a^4} - \frac{m^2\dot{a}^4}{16w^5a^4} - \frac{7m^6\dot{a}^2\ddot{a}}{8w^9a^3} + \frac{5m^4\ddot{a}\dot{a}^2}{4w^7a^3} \\
&- \frac{3m^2\dot{a}^2\ddot{a}}{8w^5a^3} - \frac{m^4\ddot{a}^2}{16w^7a^2} + \frac{m^2\ddot{a}^2}{16w^5a^2} + \frac{m^4\dot{a}\dddot{a}}{8w^7a^2} - \frac{m^2\dot{a}\dddot{a}}{8w^5a^2} .
\end{aligned}
\tag{3.51}
$$

The adiabatic renormalization subtraction terms are then defined as (we proceed in parallel to the case of scalar fields [152] [95] [96] [48, 157])

$$\langle T_{00}\rangle_{Ad} \equiv \frac{1}{2\pi^2a^3}\int_0^\infty dk k^2(\rho_k^{(0)} + \rho_k^{(2)} + \rho_k^{(4)}) . \tag{3.52}$$

Hence the renormalized 00-component of the stress-energy tensor is

$$\langle T_{00}\rangle_{ren} \equiv \langle T_{00}\rangle - \langle T_{00}\rangle_{Ad} = \frac{1}{2\pi^2a^3}\int_0^\infty dk k^2(\rho_k - \rho_k^{(0)} - \rho_k^{(2)} - \rho_k^{(4)}) . \tag{3.53}$$

This quantity is finite.

However, looking at expressions (3.50) and (3.51), one can observe that if we had subtracted only the terms up to second order, the tensor would already be convergent. In other words, the integral of the fourth-order adiabatic subtraction is, by itself, finite and independent of the mass of the field

$$-\frac{1}{2\pi^2a^3}\int_0^\infty dk k^2\rho_k^{(4)} = \frac{2}{2880\pi^2}\left[-\frac{21}{2}\frac{\dot{a}^4}{a^4} + 18\frac{\dot{a}}{a}\frac{\dddot{a}}{a} - 9\frac{\ddot{a}^2}{a^2} + 18\frac{\dot{a}^2}{a^2}\frac{\ddot{a}}{a}\right] . \tag{3.54}$$

Note that this also happens in the renormalization of a scalar field with conformal coupling $\xi = 1/6$. However, one must subtract up to the order necessary to remove the divergences for arbitrary values of ξ, and also for general metrics [157]. For an arbitrary spacetime, the fourth adiabatic order contains real divergences, which disappear accidentally for FLRW metrics in the case of fermions or scalars with $\xi = 1/6$ [60]. Therefore, according to the general rule, we subtract up to fourth adiabatic order. Discarding the fourth adiabatic subtraction would lead to a vanishing trace anomaly (see Eq. (3.74) below).

Renormalized pressure

We can also derive the vacuum expectation value of the ii-component of the Dirac stress-energy tensor (3.39) by direct computation using (3.40). Using (3.13), one should arrive at the following expression

$$\langle\bar{\psi}\gamma^i\partial_i\psi\rangle = \frac{i}{(2\pi)^3a^3}\int d^3\vec{k}\sum_{\lambda=\pm1} k_i\left[h_k^I h_k^{II*} + h_k^{I*}h_k^{II}\right]\lambda(\xi_{-\lambda}^\dagger\sigma^i\xi_{-\lambda}) . \tag{3.55}$$

The property of isotropy of the FLRW spacetime allows us to perform the calculation for $i=3$ without loss of generality. Therefore,

$$\langle\bar{\psi}\gamma^i\partial_i\psi\rangle = \frac{i2\pi}{(2\pi)^3a^3}\int_{-1}^{1}d(\cos\theta)\cos\theta\sum_{\lambda=\pm1}\lambda(\xi^\dagger_{-\lambda}\sigma^3\xi_{-\lambda})\int_0^\infty dkk^3\left[h^I_kh^{II*}_k+h^{I*}_kh^{II}_k\right], \quad (3.56)$$

where θ is the polar angle ($k_3=k\cos\theta$). Using (3.12), one finds

$$\sum_{\lambda=\pm1}\lambda(\xi^\dagger_{-\lambda}\sigma^3\xi_{-\lambda}) = -2\cos\theta \;, \quad (3.57)$$

and plugging this into (3.56), the final result for the ii-component of the stress energy tensor reads

$$\langle T_{ii}\rangle = \frac{1}{2\pi^2a}\int_0^\infty dkk^2p_k \;, \quad (3.58)$$

with

$$p_k \equiv -\frac{2k}{3a}[h^I_kh^{II*}_k+h^{I*}_kh^{II}_k] \,. \quad (3.59)$$

Again, expression (3.58) contains several ultraviolet divergences, and consequently, we must expand its integrand adiabatically and subtract from it enough terms of its expansion in order to have a finite quantity. Using the adiabatic expansion in (3.59), we get

$$p_k^{(n)} = -\frac{\omega^2-m^2}{3w}\left[FG^*+F^*G\right]^{(n)} \,. \quad (3.60)$$

The corresponding renormalized ii-component is also defined as

$$\langle T_{ii}\rangle_{ren} \equiv \langle T_{ii}\rangle - \langle T_{ii}\rangle_{Ad} = \frac{1}{2\pi^2a}\int_0^\infty dkk^2\left[p_k-p_k^{(0)}-p_k^{(2)}-p_k^{(4)}\right] \;, \quad (3.61)$$

with

$$\langle T_{ii}\rangle_{Ad} \equiv \frac{1}{2\pi^2a}\int_0^\infty dkk^2\left[p_k^{(0)}+p_k^{(2)}+p_k^{(4)}\right] \;, \quad (3.62)$$

and $[p_k^{(1)}=p_k^{(3)}=0]$

$$p_k^{(0)} = -\frac{2}{3}\left[\omega-\frac{m^2}{w}\right], \quad (3.63)$$

$$p_k^{(2)} = -\frac{m^2\dot{a}^2}{12w^3a^2}-\frac{m^2\ddot{a}}{6w^3a}+\frac{m^4\ddot{a}}{6w^5a}+\frac{m^4\dot{a}^2}{2w^5a^2}-\frac{5m^6\dot{a}^2}{12w^7a^2}, \quad (3.64)$$

$$\begin{aligned} p_k^{(4)} = {} & \frac{385m^{10}\dot{a}^4}{64w^{13}a^4}-\frac{791m^8\dot{a}^4}{64w^{11}a^4}+\frac{1477m^6\dot{a}^4}{192w^9a^4}-\frac{263m^4\dot{a}^4}{192w^7a^4}+\frac{m^2\dot{a}^4}{48w^5a^4} \\ & -\frac{77m^8\dot{a}^2\ddot{a}}{16w^{11}a^3}+\frac{77m^6\dot{a}^2\ddot{a}}{16w^9a^3}+\frac{175m^6\dot{a}^2\ddot{a}}{48w^9a^3}-\frac{175m^4\dot{a}^2\ddot{a}}{48w^7a^3}-\frac{m^4\dot{a}^2\ddot{a}}{3w^7a^3}+\frac{m^2\dot{a}^2\ddot{a}}{3w^5a^3} \\ & +\frac{7m^6\ddot{a}^2}{16w^9a^2}-\frac{5m^4\ddot{a}^2}{8w^7a^2}+\frac{3m^2\ddot{a}^2}{16w^5a^2}+\frac{7m^6\dot{a}\dddot{a}}{12w^9a^2}-\frac{5m^4\dot{a}\dddot{a}}{6w^7a^2}+\frac{m^2\dot{a}\dddot{a}}{4w^5a^2} \\ & -\frac{m^4\ddddot{a}}{24w^7a}+\frac{m^2\ddddot{a}}{24w^5a} \,. \end{aligned} \quad (3.65)$$

We also note that the integral of the fourth-order subtraction terms is finite and mass-independent

$$-\frac{1}{2\pi^2 a}\int_0^\infty dk k^2 p_k^{(4)} = \frac{2a^2}{2880\pi^2}\left[-\frac{7}{2}\frac{\dot{a}^4}{a^4} - 12\frac{\dot{a}}{a}\frac{\dddot{a}}{a} - 9\frac{\ddot{a}^2}{a^2} + 14\frac{\dot{a}^2}{a^2}\frac{\ddot{a}}{a} - 6\frac{\ddddot{a}}{a}\right] . \tag{3.66}$$

As a final comment, we stress that combining properly equations (3.9), the following simple relationship between the pressure, the energy density and the mode functions can be found

$$\rho_k = 3p_k - 2m\left[|h_k^I|^2 - |h_k^{II}|^2\right] , \tag{3.67}$$

where the second term in the right hand side is basically $\langle T_\mu^\mu \rangle_k$, with

$$\langle T_\mu^\mu \rangle = \frac{1}{2\pi^2 a^3}\int_0^\infty dk k^2 \langle T_\mu^\mu \rangle_k \ . \tag{3.68}$$

To see this, just remember that the trace is $\langle T_\mu^\mu \rangle = \langle T_{00} \rangle - \frac{3}{a^2}\langle T_{ii} \rangle$, and then, $\langle T_\mu^\mu \rangle_k = \rho_k - 3p_k$.

Expression for the renormalized stress-energy tensor

The fourth-order adiabatic subtraction terms, (3.54) and (3.66), decouple from the remaining contributions and give rise, by themselves, to a finite geometric conserved tensor. Using again the expressions in Appendix B for the different geometric quantities of a FLRW spacetime in terms of the scale factor, this conserved tensor turns out to be

$$\langle T_{\mu\nu} \rangle_{Ad}^{(4)} = \frac{2}{2880\pi^2}\left[-\frac{1}{2}{}^{(1)}H_{\mu\nu} + \frac{11}{2}{}^{(3)}H_{\mu\nu}\right] , \tag{3.69}$$

where

$$\begin{aligned} {}^{(1)}H_{\mu\nu} &= 2R_{;\mu\nu} - 2\Box R g_{\mu\nu} + 2RR_{\mu\nu} - \frac{1}{2}R^2 g_{\mu\nu} , &(3.70)\\ {}^{(3)}H_{\mu\nu} &= R_\mu^\rho R_{\rho\nu} - \frac{2}{3}RR_{\mu\nu} - \frac{1}{2}R_{\rho\sigma}R^{\rho\sigma} g_{\mu\nu} + \frac{1}{4}R^2 g_{\mu\nu} . &(3.71)\end{aligned}$$

Therefore, we get the following expression for the renormalized energy density and pressure

$$\begin{aligned}\langle T_{00} \rangle_{ren} = \frac{1}{2\pi^2 a^3}\int_0^\infty dk k^2 &\left[i\left(h_k^I \frac{\partial h_k^{I*}}{\partial t} + h_k^{II}\frac{\partial h_k^{II*}}{\partial t} - h_k^{I*}\frac{\partial h_k^I}{\partial t} - h_k^{II*}\frac{\partial h_k^{II}}{\partial t}\right)\right.\\ &\left. +2\omega + \frac{m^4 \dot{a}^2}{4\omega^5 a^2} - \frac{m^2\dot{a}^2}{4\omega^3 a^2}\right] + \langle T_{00} \rangle_{Ad}^{(4)} , \end{aligned} \tag{3.72}$$

and

$$\langle T_{ii}\rangle_{ren} = \frac{-1}{2\pi^2 a}\int_0^\infty dk \quad k^2\left[\frac{2k}{3a}[h_k^I h_k^{II*} + h_k^{I*} h_k^{II}] - \frac{2}{3}\left[\omega - \frac{m^2}{w}\right]\right. \tag{3.73}$$
$$\left. - \quad \frac{m^2\dot{a}^2}{12w^3a^2} - \frac{m^2\ddot{a}}{6w^3 a} + \frac{m^4\ddot{a}}{6w^5 a} + \frac{m^4\dot{a}^2}{2w^5a^2} - \frac{5m^6\dot{a}^2}{12w^7a^2}\right] + \langle T_{ii}\rangle_{Ad}^{(4)} ,$$

where the functions (h_k^I, h_k^{II}) above are exact solutions to the equations (3.9) and provide the mode functions defining the quantum state. Using (3.72) and (3.73) with (3.67), it is easy to see that, in the massless limit, the trace of the above tensor turns out to be

$$\langle T_\mu^\mu\rangle_{ren} = \langle T_\mu^\mu\rangle_{Ad}^{(4)} = \frac{2}{2880\pi^2}\left[-\frac{11}{2}\left(R_{\mu\nu}R^{\mu\nu} - \frac{1}{3}R^2\right) + 3\Box R\right] , \tag{3.74}$$

in exact agreement with the conformal anomaly computed by other renormalization procedures.

Before seeing some examples of this formalism, we would like to discuss briefly the interpretation of the subtraction terms in terms of redefinitions of constants in the gravitational action, and the potential ambiguities of the renormalization algorithm. As we have seen, the integrals of the zeroth and second order adiabatic subtractions in (3.52) and (3.62) do contain divergences. Following the procedure of [53], we can isolate them using dimensional regularization (n is the space-time dimension). We obtain

$$-\frac{1}{2\pi^2a^3}\int_0^\infty dk k^2 \rho_k^{(0)} \quad\to\quad -\frac{1}{2\pi^2a^3}\int_0^\infty dk k^{n-2}\rho_k^{(0)} = \frac{m^4}{8\pi^2}\frac{1}{n-4} + O(n-4) , \tag{3.75}$$
$$-\frac{1}{2\pi^2a^3}\int_0^\infty dk k^2 \rho_k^{(2)} \quad\to\quad -\frac{1}{2\pi^2a^3}\int_0^\infty dk k^{n-2}\rho_k^{(2)} = \frac{m^2}{8\pi^2}\frac{1}{n-4}\frac{\dot{a}^2}{a^2} + O(n-4) , \tag{3.76}$$

and

$$-\frac{1}{2\pi^2a}\int_0^\infty dk k^2 p_k^{(0)} \quad\to\quad -\frac{1}{2\pi^2a}\int_0^\infty dk k^{n-2} p_k^{(0)} = -\frac{m^4a^2}{8\pi^2}\frac{1}{n-4} + O(n-4) , \tag{3.77}$$
$$-\frac{1}{2\pi^2a}\int_0^\infty dk k^2 p_k^{(2)} \quad\to\quad -\frac{1}{2\pi^2a}\int_0^\infty dk k^{n-2} p_k^{(2)} = -\frac{m^2a^2}{24\pi^2}\frac{1}{n-4}\left(2\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2}\right) + O(n-4) \tag{3.78}$$

The zeroth and second order adiabatic subtraction terms, (3.75)-(3.76) and (3.77)-(3.78), are formally divergent, but can be suitable expressed as geometric tensors. Using the geometric identities of Appendix B one can easily find

$$\langle T_{\mu\nu}\rangle_{Ad}^{(0)} = \frac{m^4}{8\pi^2}\frac{1}{n-4}g_{\mu\nu} + O(n-4) , \tag{3.79}$$
$$\langle T_{\mu\nu}\rangle_{Ad}^{(2)} = -\frac{m^2}{24\pi^2}\frac{1}{n-4}G_{\mu\nu} + O(n-4) , \tag{3.80}$$

with $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$, the Einstein tensor.

Recall now the Einstein's gravitational field equations

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = -8\pi G \langle T_{\mu\nu} \rangle \ . \tag{3.81}$$

Equation (3.79) suggests the possibility of absorbing the UV-divergence of the zeroth adiabatic order into the cosmological constant, Λ, while expression (3.80) offers the possibility of renormalizing the second adiabatic order divergence into the Newton's universal constant, G. This way, the adiabatic subtraction terms can be nicely interpreted in terms of renormalization of coupling constants, in parallel to the scalar fields case [53].

On the other hand, as we have stressed before, in a general space-time the fourth-order subtraction terms give rise to proper UV divergencies [60]. They turn out to be proportional to a linear combination of the two independent geometric tensors with the appropriate dimensions, namely ${}^{(1)}H_{\mu\nu}$ and ${}^{(2)}H_{\mu\nu}$. The four type of divergent subtraction terms, proportional to $m^4 g_{\mu\nu}, m^2 G_{\mu\nu}$, ${}^{(1)}H_{\mu\nu}$ and ${}^{(2)}H_{\mu\nu}$, generate intrinsic ambiguities in the curved space renormalization program for the stress-energy tensor [94, 185]. The first two can be naturally associated to the renormalization of the cosmological constant and Newton's constant. In our FLRW spacetime ${}^{(2)}H_{\mu\nu}$ is proportional to ${}^{(1)}H_{\mu\nu}$, and hence we are left with only one relevant renormalization parameter ambiguity. This translates to the fact that the general expression for the finite fourth order adiabatic contribution $\langle T_{\mu\nu} \rangle^{(4)}_{Ad}$ contains actually an arbitrary coefficient c_1

$$\langle T_{\mu\nu} \rangle^{(4)}_{Ad} = \frac{2}{2880\pi^2} \left[c_1 \, {}^{(1)}H_{\mu\nu} + \frac{11}{2} \, {}^{(3)}H_{\mu\nu} \right] \ . \tag{3.82}$$

Our adiabatic regularization methods leads to $c_1 = -1/2$. Other renormalization methods can only potentially differ from our results on the value of this coefficient. However, we remark that the ambiguity disappears for spacetime backgrounds for which the tensor ${}^{(1)}H_{\mu\nu}$ vanishes. This happens for physically relevant space-times, like de Sitter space or the radiation-dominated universe.

Note that if one considers the general expression (3.82), instead of (3.69), the numerical coefficient of $\Box R$ in (3.74) is actually proportional to c_1.

Stress-energy conservation

The above $\langle T_{\mu\nu} \rangle_{ren}$ is, as expected, a conserved tensor $\nabla^\mu \langle T_{\mu\nu} \rangle_{ren} = 0$. The conservation equations in a FLRW spacetime can be spelled out as

$$\langle T^{0\nu}{}_{;\nu} \rangle = \langle \dot{T}_{00} \rangle + 3\frac{\dot{a}}{a} \langle T_{00} \rangle + \frac{3}{a^2}\frac{\dot{a}}{a} \langle T_{ii} \rangle = 0 \ , \tag{3.83}$$

$$\langle T^{i\nu}{}_{;\nu} \rangle = 0, \qquad i = 1,2,3 \tag{3.84}$$

and they can be checked by direct computation. This is a consequence of the fact that, for each adiabatic order of the formal subtraction tensors $\langle T_{\mu\nu}\rangle^{(n)}_{Ad}$, where

$$\langle T_{00}\rangle^{(n)}_{Ad} = -\frac{1}{2\pi^2 a^3}\int_0^\infty dk k^2 \rho^{(n)}_k \ , \tag{3.85}$$

$$\langle T_{ii}\rangle^{(n)}_{Ad} = -\frac{1}{2\pi^2 a}\int_0^\infty dk k^2 p^{(n)}_k \ , \tag{3.86}$$

we have the independent conservation laws $\nabla^\mu \langle T_{\mu\nu}\rangle^{(n)}_{Ad} = 0$, for $n = 0,2,4$.

3.4 Examples

In this section, we work out the renormalized stress-energy tensor for two different spacetimes: de Sitter space-time and a radiation-dominated universe. For a given scale factor $a(t)$, we need to solve (3.9). As it is a system of two coupled first-order differential equations, we have two free parameters which have to be fixed somehow. More specifically, given a particular solution (h^I_k, h^{II}_k) normalized as in (3.15), we can construct the general solution by a Bogolubov-type rotation

$$\begin{aligned} h^I_k &\rightarrow E_k h^I_k + F_k h^{II*}_k \ , \\ h^{II}_k &\rightarrow E_k h^{II}_k - F_k h^{I*}_k \ . \end{aligned} \tag{3.87}$$

where E_k and F_k are two arbitrary complex-valued constants. On the other hand, we should also ensure the normalization condition (3.15), which implies the following constraint

$$|E_k|^2 + |F_k|^2 = 1 \ . \tag{3.88}$$

Note that the renormalization ambiguity associated to the ${}^{(1)}H_{\mu\nu}$ tensor disappears for de Sitter space-time and the radiation-dominated universe.

Renormalized stress-energy tensor in de Sitter spacetime

For de Sitter spacetime $a(t) = e^{Ht}$ with H a constant, the general solution to the field equations (3.10) and (3.11) can be conveniently expressed, using the transformation in (3.87), as the following linear combination

$$h^I_k(t) = E_k\left(\frac{i}{2}\sqrt{\pi z}\, e^{\frac{\pi\mu}{2}} H^{(1)}_{\frac{1}{2}-i\mu}(z)\right) + F_k\left(\frac{1}{2}\sqrt{\pi z}\, e^{\frac{\pi\mu}{2}} H^{(1)}_{-\frac{1}{2}-i\mu}(z)\right)^* \ , \tag{3.89}$$

$$h^{II}_k(t) = E_k\left(\frac{1}{2}\sqrt{\pi z}\, e^{\frac{\pi\mu}{2}} H^{(1)}_{-\frac{1}{2}-i\mu}(z)\right) - F_k\left(\frac{i}{2}\sqrt{\pi z}\, e^{\frac{\pi\mu}{2}} H^{(1)}_{\frac{1}{2}-i\mu}(z)\right)^* \ . \tag{3.90}$$

where $z \equiv kH^{-1}e^{-Ht}$, $\mu \equiv m/H$, $H^{(1)}(z)$ are Hankel functions of the first kind and E_k and F_k are constants that need to be fixed with appropriate initial conditions. A crucial physical requirement is that, as $k \to \infty$, the physical solutions should have the adiabatic asymptotic form

$$h_k^I \sim \sqrt{\frac{\omega+m}{2\omega}}\, e^{-i\int^{t'}\omega(t')dt'} , \qquad h_k^{II} \sim \sqrt{\frac{\omega-m}{2\omega}}\, e^{-i\int^{t'}\omega(t')dt'} . \tag{3.91}$$

This way, one recovers in this limit the Minkowskian solutions. This leads to

$$E_k \sim 1 , \qquad F_k \sim 0 , \tag{3.92}$$

as $k \to \infty$. The above condition can be naturally achieved by the simple solution $E_k = 1$ and $F_k = 0$. This determines a vacuum for spin one-half fields analogous to the Bunch-Davies vacuum [171] for scalar fields. It is also the natural extension of the conformal vacuum for massless fields. It can be uniquely characterized by invoking de Sitter invariance [157].

By changing the integration variable from k to z, we obtain that the energy density (3.53) is

$$\langle T_{00}\rangle_{ren} = \frac{H^3}{2\pi^2}\int_0^\infty dz z^2(\rho_k - \rho_k^{(0)} - \rho_k^{(2)} - \rho_k^{(4)}) , \tag{3.93}$$

where from (3.89), (3.90) and (3.92), the bare contribution (3.42) is

$$\rho_k(z) = \frac{\pi H e^{\pi\mu} z^2}{2}\left(\frac{\mu}{z}\left[H^{(1)}_{\nu-1}H^{(2)}_{-\nu} - H^{(1)}_{\nu}H^{(2)}_{1-\nu}\right] + i\left[H^{(1)}_{\nu-1}H^{(2)}_{1-\nu} - H^{(1)}_{\nu}H^{(2)}_{-\nu}\right]\right) , \tag{3.94}$$

with $\nu \equiv (1/2) - i\mu$ and $H^{(1,2)}_\nu = H^{(1,2)}_\nu(z)$. On the other hand, the subtraction terms, from (3.45), (3.50) and (3.51), take the form

$$\rho_k^{(0)} = -2H\sqrt{z^2+\mu^2} , \tag{3.95}$$

$$\rho_k^{(2)} = H\left(\frac{\mu^2}{4(z^2+\mu^2)^{3/2}} - \frac{\mu^4}{4(z^2+\mu^2)^{5/2}}\right) , \tag{3.96}$$

$$\rho_k^{(4)} = H\left(\frac{105\mu^8}{64(z^2+\mu^2)^{11/2}} - \frac{119\mu^6}{32(z^2+\mu^2)^{9/2}} + \frac{165\mu^4}{64(z^2+\mu^2)^{7/2}} - \frac{\mu^2}{2(z^2+\mu^2)^{5/2}}\right) . \tag{3.97}$$

Similar expressions can be obtained for the renormalized pressure (3.61),

$$\langle T_{ii}\rangle_{ren} = \frac{a^2H^3}{2\pi^2}\int_0^\infty dz z^2(p_k - p_k^{(0)} - p_k^{(2)} - p_k^{(4)}) , \tag{3.98}$$

where, from (3.59),

$$p_k(z) = i\frac{\pi H e^{\pi\mu} z^2}{6}\left[H^{(1)}_{\nu-1}H^{(2)}_{1-\nu} - H^{(1)}_{\nu}H^{(2)}_{-\nu}\right] . \tag{3.99}$$

From these results, one can obtain $\langle T_{00}\rangle_{ren}$ and $\langle T_{ii}\rangle_{ren}$ numerically with very high accuracy. We reproduce exactly the analytical expression already obtained in [126, 127] from the trace anomaly and the symmetries of de Sitter space-time

$$\langle T_{\mu\nu}\rangle_r = \frac{1}{960\pi^2} g_{\mu\nu}\left[11H^4 + 130H^2m^2 + 120m^2(H^2+m^2)\left(\log\left(\frac{m}{H}\right) - \mathfrak{Re}\left[\psi\left(-1+i\frac{m}{H}\right)\right]\right)\right] , \tag{3.100}$$

where $\psi(z)$ is the digamma function.

Radiation dominated Universe

In this section, we apply our general results for a radiation dominated universe. This is also a nice example to show how the general procedure works. In this case, the two independent solutions of the differential equation for the field modes with $a(t) = a_0\sqrt{t}$ are given in terms of the Whittaker functions $\frac{1}{a(t)}W_{\kappa,\mu}(z)$ and $\frac{1}{a(t)}W_{-\kappa,\mu}(-z)$ (see [40]), where

$$\kappa = \frac{1}{4} - ix^2, \quad \mu = \frac{1}{4}, \quad z = i2mt , \tag{3.101}$$

with $x^2 \equiv k^2/(a_0^2 2m)$. We choose a set of two linear independent solutions for the field modes of the form

$$h_k^I = E_k\left(N\frac{W_{\kappa,\mu}(z)}{\sqrt{a(t)}}\right) + F_k\left(N\frac{k}{2ma(t)^{3/2}}\left[W_{\kappa,\mu}(z) + \left(\kappa - \frac{3}{4}\right)W_{\kappa-1,\mu}(z)\right]\right)^* , \tag{3.102}$$

where the constant $N = \frac{a_0^{1/2}}{(2m)^{1/4}}e^{-\frac{\pi}{2}x^2}$ and the condition $|E_k|^2 + |F_k|^2 = 1$ are fixed from the normalization condition (3.15). The adiabatic condition (3.91) for $k \to \infty$ also requires that $E_k \sim 1$ and $F_k \sim 0$. Moreover, a detailed analysis (see Appendix C) of the asymptotic properties of the stress energy tensor components using the Whittaker functions [130] allows us to characterize the condition for the renormalizability of the vacuum expectation values of the stress-tensor as

$$|E_k|^2 - |F_k|^2 = 1 + O(k^{-5}) . \tag{3.103}$$

As one can see from Appendix C, this particular combination of E_k and F_k is the crucial one in the analysis of the renormalized energy density (3.308) and pressure (3.309) for 1/2 spin fields.

In contrast with the previous example of de Sitter space-time, the absence of extra symmetries (in addition to the standard homogeneity and isotropy of a FLRW space-time) for the radiation-dominated background does not allow us to fix a natural preferred vacuum state. However, the early and late-time behaviors ($t << m^{-1}$ and $t >> m^{-1}$, respectively) of the renormalized stress-energy tensor can be obtained generically, and agree with the forms assumed by classical cosmology. As detailed in Appendix C, we have that, as time evolves and reaches the regime $t >> m^{-1}$, the renormalized energy density takes the form of cold matter

$$\langle T_{00}\rangle_{ren} \sim \frac{\rho_{0m}}{a^3} \ , \tag{3.104}$$

where

$$\rho_{0m} = \frac{m}{\pi^2}\int_0^\infty dk k^2 \left[1-(|E_k|^2-|F_k|^2)\right] \geq 0 \ . \tag{3.105}$$

Notice that, $2 \geq 1-(|E_k|^2-|F_k|^2) = 2|F_k|^2 \geq 0$, and together with the renormalizability condition (3.103), we see that the energy density ρ_{0m} is finite and definite positive. The specific value of ρ_{0m} depends on the form of the quantum state for our spin-1/2 field, i.e. of the choice of E_k and F_k. Since at late times $t >> m^{-1}$ the relation (3.102) transforms into

$$h_k^I(t) \sim E_k\sqrt{\frac{\omega+m}{2\omega}}\, e^{-i\int^{t'}\omega(t')dt'} + F_k\sqrt{\frac{\omega-m}{2\omega}}\, e^{-i\int^{t'}\omega(t')dt'} \ , \tag{3.106}$$

the coefficients F_k are actually the fermionic β-type (Bogolubov) coefficients [147, 148]. Therefore, we actually get $\rho_{0m} \sim m\langle n(t)\rangle$, where $\langle n(t)\rangle$ is the number density of the created particles. Moreover, we find

$$\frac{\langle T_{ii}\rangle_{ren}}{a^2} \sim 0 \ , \tag{3.107}$$

and hence the pressure obeys the cold matter equation of state.

On the other hand, for sufficiently early times in the evolution, $t << m^{-1}$, we have (see Appendix C)

$$\langle T_{00}\rangle_{ren} \sim \frac{\rho_{0r}}{a^4} \ , \tag{3.108}$$

with

$$\rho_{0r} = \frac{1}{\pi^2}\int_0^\infty dk k^3 \left[1-(|E_k|^2-|F_k|^2)\right] \geq 0 \ , \tag{3.109}$$

and additionally

$$\frac{\langle T_{ii}\rangle_{ren}}{a^2} \sim \frac{1}{3}\langle T_{00}\rangle_{ren} \ . \tag{3.110}$$

Note again that ρ_{0r} is finite and definite positive. The specific value of ρ_{0r} depends on the specific form of the quantum state throughout the complex functions E_k and F_k. From (3.110), we see that $p \sim \rho/3$, in agreement with the assumptions of classical cosmology for the radiation.

The analysis and phenomenology of the renormalized stress-energy tensor obtained from specific choices of the vacuum state is beyond the scope of the present work. Note again that any choice for the quantum state has not a preferred status, in contrast with the Bunch-Davies type vacuum of the previous example, due to the absence of the additional symmetries endowed by de Sitter space-time.

3.5 Quantized Dirac field with Yukawa coupling. Einstein's semiclassical equations

We consider now a theory defined by the action functional $S = S[g_{\mu\nu}, \Phi, \psi, \nabla\psi]$, where ψ represents a Dirac field, Φ is a scalar field, and $g_{\mu\nu}$ stands for the spacetime metric. We decompose the action as $S = S_g + S_m$, where S_m is the matter sector

$$S_m = \int d^4x \sqrt{-g}\ \left\{ \frac{i}{2}[\bar{\psi}\underline{\gamma}^{\mu}\nabla_{\mu}\psi - (\nabla_{\mu}\bar{\psi})\underline{\gamma}^{\mu}\psi)] - m\bar{\psi}\psi - g_Y\Phi\bar{\psi}\psi \right\} , \tag{3.111}$$

and S_g is the gravity-scalar sector, which will be presented in the next subsection. Here, $\underline{\gamma}^{\mu}(x)$ are the spacetime-dependent Dirac matrices satisfying the anticommutation relations $\{\underline{\gamma}^{\mu}, \underline{\gamma}^{\nu}\} = 2g^{\mu\nu}$, related to the usual Minkowski ones by the vierbein field $V_{\mu}^{a}(x)$, defined through $g_{\mu\nu}(x) = V_{\mu}^{a}(x)V_{\nu}^{b}(x)\eta_{ab}$. On the other hand, ∇_{μ} is the Levi-Civita connection, which in a suitable coordinate chart we can express as $\nabla_{\mu} \equiv \partial_{\mu} - \Gamma_{\mu}$ with Γ_{μ} the 1-form spin-connection; m is the mass of the Dirac field; and g_Y is the dimensionless coupling constant of the Yukawa interaction. In (3.111), both the metric $g_{\mu\nu}(x)$ and the scalar field $\Phi(x)$ are regarded as classical external fields. The Dirac spinor $\psi(x)$ will be our quantized field, living in a curved spacetime and possessing a Yukawa coupling to the classical field Φ. The Dirac equation is

$$(i\underline{\gamma}^{\mu}\nabla_{\mu} - m - g_Y\Phi)\psi = 0 , \tag{3.112}$$

and the stress-energy tensor is given by [48]

$$T_{\mu\nu}^{m} := \frac{2}{\sqrt{-g}}\frac{\delta S_m}{\delta g^{\mu\nu}} = \frac{V_{\nu a}}{\det V}\frac{\delta S_m}{\delta V_a^{\mu}} = \frac{i}{2}\left[\bar{\psi}\underline{\gamma}_{(\mu}\nabla_{\nu)}\psi - (\nabla_{(\mu}\bar{\psi})\underline{\gamma}_{\nu)}\psi\right] . \tag{3.113}$$

The presence of the Yukawa interaction with the external field Φ modifies the standard conservation equation. We have, instead,

$$\nabla_{\mu}T_{m}^{\mu\nu} = g_Y\bar{\psi}\psi\nabla^{\nu}\Phi . \tag{3.114}$$

These equations can be easily seen as the consequence of the invariance of the action functional S under spacetime diffeomorphisms $\delta x^\mu = \epsilon^\mu(x)$: $\delta\Phi = \epsilon^\mu \nabla_\mu \Phi$, $\delta g_{\mu\nu} = 2\nabla_{(\mu}\epsilon_{\nu)}$. One gets

$$\nabla_\mu T_m^{\mu\nu} + \frac{1}{\sqrt{-g}}\frac{\delta S_m}{\delta\Phi}\nabla^\nu\Phi = 0 \ , \tag{3.115}$$

which reproduces (3.114). We will assume that the quantum theory fully respects this symmetry. Therefore, we demand

$$\langle \nabla_\mu T_m^{\mu\nu}\rangle = g_Y \langle\bar\psi\psi\rangle \nabla^\nu\Phi \ . \tag{3.116}$$

Adding the gravity-scalar sector

The complete theory, including the gravity-scalar sector in the action, can be described by

$$S = S_g + S_m = \frac{1}{16\pi G}\int d^4x\sqrt{-g}R(g) + \int d^4x\sqrt{-g}\left\{\frac{1}{2}g^{\mu\nu}\nabla_\mu\Phi\nabla_\nu\Phi - V(\Phi)\right\} + S_m \ , \tag{3.117}$$

where S_m is the action for the matter sector given in (3.111). We will reconsider the form of the action in Section 3.8, in view of the counterterms required to cancel the UV divergences of the quantized Dirac field. However, let us work for the moment with the action (3.117). The Einstein equations are then

$$G^{\mu\nu} + 8\pi G(\nabla^\mu\Phi\nabla^\nu\Phi - \frac{1}{2}g^{\mu\nu}\nabla^\rho\Phi\nabla_\rho\Phi + g^{\mu\nu}V(\Phi)) = -8\pi G T_m^{\mu\nu} \ , \tag{3.118}$$

and the equation for the scalar field is

$$\Box\Phi + \frac{\partial V}{\partial\Phi} = -g_Y\bar\psi\psi \ . \tag{3.119}$$

The semiclassical equations are obtained from (3.118) and (3.119) by replacing $T_m^{\mu\nu}$ and $\bar\psi\psi$ by the corresponding (renormalized) vacuum expectation values $\langle T_m^{\mu\nu}\rangle_{ren}$ and $\langle\bar\psi\psi\rangle_{ren}$,

$$G^{\mu\nu} + 8\pi G(\nabla^\mu\Phi\nabla^\nu\Phi - \frac{1}{2}g^{\mu\nu}\nabla^\rho\Phi\nabla_\rho\Phi + g^{\mu\nu}V(\Phi)) = -8\pi G\langle T_m^{\mu\nu}\rangle_{ren} \ , \tag{3.120}$$

$$\Box\Phi + \frac{\partial V}{\partial\Phi} = -g_Y\langle\bar\psi\psi\rangle_{ren} \ . \tag{3.121}$$

These equations are consistent with the Bianchi identities $\nabla_\mu G^{\mu\nu} = 0$, since

$$\nabla_\mu(\nabla^\mu\Phi\nabla^\nu\Phi - \frac{1}{2}g^{\mu\nu}\nabla^\rho\Phi\nabla_\rho\Phi + g^{\mu\nu}V(\Phi)) = (\Box\Phi + \frac{\partial V}{\partial\Phi})\nabla^\nu\Phi \ , \tag{3.122}$$

and, from (3.116) and (3.121), we have

$$\nabla_\mu \langle T_m^{\mu\nu}\rangle_{ren} = g_Y \langle\bar{\psi}\psi\rangle_{ren}\nabla^\nu\Phi = -(\Box\Phi + \frac{\partial V}{\partial\Phi})\nabla^\nu\Phi \ . \quad (3.123)$$

When the spacetime is an expanding universe ($ds^2 = dt^2 - a^2(t)d\vec{x}^2$), and Φ is an homogeneous scalar field $\Phi = \Phi(t)$ (e.g. an inflaton), the equations (3.120) and (3.121) describe the backreaction on the metric-inflaton system due to matter particle production and vacuum polarization, codified in the renormalized vacuum expectation values $\langle T_m^{\mu\nu}\rangle_{ren}$ and $\langle\bar{\psi}\psi\rangle_{ren}$. It is then important to elaborate an efficient method to compute these quantities in this cosmological setting.

3.6 Adiabatic expansion for a Dirac field with Yukawa coupling

Recall that in a spatially flat FLRW spacetime, the time-dependent gamma matrices are related with the Minkowskian ones by $\underline{\gamma}^0(t) = \gamma^0$ and $\underline{\gamma}^i(t) = \frac{\gamma^i}{a(t)}$, and the components of the spin-connections are $\Gamma_0 = 0$ and $\Gamma_i = \frac{\dot{a}}{2}\gamma_0\gamma_i$. The Dirac equation with the Yukawa interaction $i\underline{\gamma}^\mu\nabla_\mu\psi - m\psi = g_Y\Phi\psi$, taking Φ as a homogenous scalar field $\Phi = \Phi(t)$, can be written in the FLRW coordinate chart as

$$\left(\partial_0 + \frac{3}{2}\frac{\dot{a}}{a} + \frac{1}{a}\gamma^0\vec{\gamma}\vec{\nabla} + i(m + s(t))\gamma^0\right)\psi = 0 \ , \quad (3.124)$$

where we have defined $s(t) \equiv g_Y\Phi(t)$. If we expand the field ψ as $\psi = \int \frac{d^3\vec{k}}{(2\pi)^{3/2}}\psi_{\vec{k}}(t)e^{i\vec{k}\vec{x}}$, and we substitute it into (3.124), we obtain the following differential equation for $\psi_{\vec{k}}$:

$$\left(\partial_t + \frac{3\dot{a}}{2a} + i\gamma^0\vec{\gamma}\frac{\vec{k}}{a} + i\gamma^0(m + s(t))\right)\psi_{\vec{k}} = 0 \ . \quad (3.125)$$

In order to solve this equation, it is convenient to write the Dirac field in terms of two two-component spinors of the generic form

$$\psi_{\vec{k},\lambda}(t) = \frac{1}{a^{3/2}(t)}\begin{pmatrix} h_k^I(t)\xi_\lambda(\vec{k}) \\ h_k^{II}(t)\frac{\vec{\sigma}\vec{k}}{k}\xi_\lambda(\vec{k}) \end{pmatrix} , \quad (3.126)$$

where ξ_λ with $\lambda = \pm 1$ are two constant orthonormal two-spinors ($\xi_\lambda^\dagger\xi_{\lambda'} = \delta_{\lambda,\lambda'}$), eigenvectors of the helicity operator $\frac{\vec{\sigma}\vec{k}}{2k}\xi_\lambda = \frac{\lambda}{2}\xi_\lambda$. The explicit forms of ξ_{+1} and ξ_{-1} are given in (3.12). where $\vec{k} = (k_1, k_2, k_3)$ and $|\vec{k}| = k$. The time-dependent functions h_k^I and h_k^{II} satisfy now the first-order coupled equations

$$h_k^{II} = \frac{ia}{k}\left(\frac{\partial h_k^I}{\partial t} + i(m + s(t))h_k^I\right) \ , \quad h_k^I = \frac{ia}{k}\left(\frac{\partial h_k^{II}}{\partial t} - i(m + s(t))h_k^{II}\right) \ . \quad (3.127)$$

Given a particular solution $\{h_k^I(t),\, h_k^{II}(t)\}$ to equations (3.127), one can construct the modes as in (3.8). This will be a solution of positive-frequency type in the adiabatic regime. A solution of negative-frequency type can be obtained by applying a charge conjugate transformation $C\psi = i\gamma^2\psi^*$ and reads as in (3.13). Recall that the Dirac inner product is defined as in (3.14): $(\psi_1,\psi_2) = \int d^3x a^3 \psi_1^\dagger \psi_2$; and that the normalization condition for the above four-spinors $(u_{\vec{k}\lambda}, v_{\vec{k}'\lambda'}) = 0$, $(u_{\vec{k}\lambda}, u_{\vec{k}'\lambda'}) = (v_{\vec{k}\lambda}, v_{\vec{k}'\lambda'}) = \delta_{\lambda\lambda'}\delta^{(3)}(\vec{k}-\vec{k}')$, reduces to equation (3.15):

$$|h_k^I|^2 + |h_k^{II}|^2 = 1 \ . \tag{3.128}$$

Since the Dirac scalar product is preserved by the cosmological evolution, the normalization condition (3.128) holds at any time. This ensures also the standard anticommutation relations for the creation and annihilation operators ($\{B_{\vec{k},\lambda}, B^\dagger_{\vec{k}',\lambda'}\} = \delta^3(\vec{k}-\vec{k}')\delta_{\lambda\lambda'}$, $\{B_{\vec{k},\lambda}, B_{\vec{k}',\lambda'}\} = 0$, and similarly for the $D_{\vec{k},\lambda}$, $D^\dagger_{\vec{k}',\lambda'}$ operators), introduced by the Fourier expansion of the Dirac field operator (3.16)

$$\psi(x) = \int d^3\vec{k} \sum_\lambda \left[B_{\vec{k}\lambda} u_{\vec{k}\lambda}(x) + D^\dagger_{\vec{k}\lambda} v_{\vec{k}\lambda}(x) \right] \ . \tag{3.129}$$

Adiabatic expansion

We now compute the adiabatic expansion of a Dirac field living in a FLRW spacetime, and possessing a Yukawa interaction term with a classical background field. We proceed similarly to previous sections. We know that, in the adiabatic limit, and in the absence of interaction, the natural solution of the field modes h_k^I and h_k^{II} is

$$h_k^I(t) \sim \sqrt{\frac{\omega(t)+m}{2\omega(t)}}\, e^{-i\int^t \omega(t')dt'} \ , \qquad h_k^{II}(t) \sim \sqrt{\frac{\omega(t)-m}{2\omega(t)}}\, e^{-i\int^t \omega(t')dt'} \ , \tag{3.130}$$

where $\omega = \sqrt{\frac{k^2}{a^2} + m^2}$ is the frequency of the field mode. This will constitute the zeroth-order term of the adiabatic expansion. Mimicking the ansatz introduced in (3.17) and in [126, 127], we write the h_k^I and h_k^{II} functions as

$$h_k^I(t) = \sqrt{\frac{\omega(t)+m}{2\omega(t)}}\, e^{-i\int^t \Omega(t')dt'} F(t) \ , \qquad h_k^{II}(t) = \sqrt{\frac{\omega(t)-m}{2\omega(t)}}\, e^{-i\int^t \Omega(t')dt'} G(t) \ , \tag{3.131}$$

where $\Omega(t)$, $F(t)$ and $G(t)$ are time-dependent functions, which we expand adiabatically as [since this case is more general than (3.18) we use the same letters, no confusion should arise]

$$\begin{aligned}\Omega &= \omega+\omega^{(1)}+\omega^{(2)}+\omega^{(3)}+\omega^{(4)}+\dots\ ,\\ F &= 1+F^{(1)}+F^{(2)}+F^{(3)}+F^{(4)}+\dots\ ,\\ G &= 1+G^{(1)}+G^{(2)}+G^{(3)}+G^{(4)}+\dots\ .\end{aligned} \tag{3.132}$$

Here, $F^{(n)}$, $G^{(n)}$ and $\omega^{(n)}$ are terms of nth adiabatic order. By substituting (3.131) into the equations of motion (3.127) and the normalization condition (3.128), we obtain the following system of three equations,

$$\begin{aligned}(\omega-m)G = \Omega F + i\dot{F} + \frac{iF}{2}\frac{d\omega}{dt}\left(\frac{1}{\omega+m}-\frac{1}{\omega}\right)-(m+s)F\ ,\\ (\omega+m)F = \Omega G + i\dot{G} + \frac{iG}{2}\frac{d\omega}{dt}\left(\frac{1}{\omega-m}-\frac{1}{\omega}\right)+(m+s)G\ ,\\ (\omega+m)FF^{*}+(\omega-m)GG^{*}=2\omega\ .\end{aligned} \tag{3.133}$$

To obtain the expressions for $\Omega^{(n)}$, $F^{(n)}$, and $G^{(n)}$, we introduce the adiabatic expansions (3.132) into (3.133), and solve order by order. As usual, we consider $\dot{a}$ of adiabatic order 1, $\ddot{a}$ of adiabatic order 2, and so on. On the other hand, we consider the interaction term $s(t)$ of adiabatic order 1, so that the zeroth order term in (3.131) recovers the free field solution in the adiabatic limit, defined in (3.130). Similarly, time-derivatives of the interaction increase the adiabatic order, so that $\dot{s}$ is of order 2, $\ddot{s}$ of order 3, and so on. With this, a generic expression $f^{(n)}$ of adiabatic order n (e.g. $f^{(n)}=F^{(n)},G^{(n)},\Omega^{(n)}$) will be written as a sum of all possible products of nth adiabatic order formed by s, a, and their time-derivatives. For example, functions of adiabatic orders 1 and 2 will be written respectively as

$$\begin{aligned}f^{(1)} &= \alpha_1 s+\alpha_2\dot{a}\ ,\\ f^{(2)} &= \beta_1 s^2+\beta_2\dot{s}+\beta_3\ddot{a}+\beta_4\dot{a}^2+\beta_5\dot{a}s\ ,\end{aligned} \tag{3.134}$$

with $\alpha_n \equiv \alpha_n(m,k,a)$ and $\beta_n \equiv \beta_n(m,k,a)$. The assignment of s as adiabatic order 1 is consistent with the scaling dimension of the scalar field, as it possesses the same dimensions as $\dot{a}$.

First adiabatic order

By keeping only terms of first adiabatic order in (3.133), the system of three equations gives

$$
\begin{aligned}
(\omega-m)G^{(1)} &= (\omega-m)F^{(1)}+\omega^{(1)}-s+\frac{i}{2}\frac{d\omega}{dt}\left(\frac{1}{\omega+m}-\frac{1}{\omega}\right) ,\\
(\omega+m)F^{(1)} &= (\omega+m)G^{(1)}+\omega^{(1)}+s+\frac{i}{2}\frac{d\omega}{dt}\left(\frac{1}{\omega-m}-\frac{1}{\omega}\right) ,\\
(\omega+m)(F^{(1)}+F^{(1)*})&+(\omega-m)(G^{(1)}+G^{(1)*})=0 . \qquad (3.135)
\end{aligned}
$$

We now treat independently the real and imaginary parts by writing $F^{(1)}=f_x^{(1)}+if_y^{(1)}$ and $G^{(1)}=g_x^{(1)}+ig_y^{(1)}$. We obtain for the real part

$$
\begin{aligned}
(\omega-m)(g_x^{(1)}-f_x^{(1)}) &= \omega^{(1)}-s ,\\
(\omega+m)(g_x^{(1)}-f_x^{(1)}) &= -\omega^{(1)}-s ,\\
(\omega+m)f_x^{(1)}+(\omega-m)g_x^{(1)} &= 0 , \qquad (3.136)
\end{aligned}
$$

which has as solutions

$$
f_x^{(1)}=\frac{s}{2\omega}-\frac{ms}{2\omega^2} , \quad g_x^{(1)}=-\frac{s}{2\omega}-\frac{ms}{2\omega^2} , \quad \omega^{(1)}=\frac{ms}{\omega} . \qquad (3.137)
$$

On the other hand, the imaginary part of the system gives

$$
\begin{aligned}
(\omega-m)(g_y^{(1)}-f_y^{(1)}) &= \frac{1}{2}\frac{d\omega}{dt}\left(\frac{1}{\omega+m}-\frac{1}{\omega}\right) ,\\
(\omega+m)(g_y^{(1)}-f_y^{(1)}) &= -\frac{1}{2}\frac{d\omega}{dt}\left(\frac{1}{\omega-m}-\frac{1}{\omega}\right) . \qquad (3.138)
\end{aligned}
$$

These two equations are not independent. The obtained solution for $g_y^{(1)}$ and $f_y^{(1)}$ is

$$
f_y^{(1)}=A-\frac{m\dot{a}}{2a\omega^2} , \quad g_y^{(1)}=A , \qquad (3.139)
$$

where A is an arbitrary first-order adiabatic function. We will choose the simplest solution

$$
f_y^{(1)}=-\frac{m\dot{a}}{4\omega^2 a} , \quad g_y^{(1)}=\frac{m\dot{a}}{4\omega^2 a} , \qquad (3.140)
$$

obeying the condition $F^{(1)}(m,s)=G^{(1)}(-m,-s)$. Therefore, the adiabatic expansion will also preserve the symmetries of the equations (3.127) with respect to the change $(m,s)\rightarrow(-m,-s)$. We have checked that physical expectation values are independent to any potential ambiguity in this kind of choice.

Second adiabatic order

In the same way, the second-order terms of (3.133) give

$$(\omega-m)G^{(2)}=(\omega-m)F^{(2)}+(\omega^{(1)}-s)F^{(1)}+\omega^{(2)}+i\dot{F}^{(1)}+i\frac{F^{(1)}}{2}\frac{d\omega}{dt}\left(\frac{1}{\omega+m}-\frac{1}{\omega}\right),$$
$$(\omega+m)F^{(2)}=(\omega+m)G^{(2)}+(\omega^{(1)}+s)G^{(1)}+\omega^{(2)}+i\dot{G}^{(1)}+i\frac{G^{(1)}}{2}\frac{d\omega}{dt}\left(\frac{1}{\omega-m}-\frac{1}{\omega}\right),$$
$$(\omega+m)(F^{(2)}+F^{(1)}F^{(1)*}+F^{(2)*})+(\omega-m)(G^{(2)}+G^{(1)}G^{(1)*}+G^{(2)*})=0\,, \quad (3.141)$$

where the first-order terms have already been deduced above. Taking the real part of these equations, we obtain

$$(\omega-m)(g_x^{(2)}-f_x^{(2)})=(\omega^{(1)}-s)f_x^{(1)}+\omega^{(2)}-\dot{f}_y^{(1)}-\frac{f_y^{(1)}}{2}\frac{d\omega}{dt}\left(\frac{1}{\omega+m}-\frac{1}{\omega}\right),$$
$$(\omega+m)(g_x^{(2)}-f_x^{(2)})=-(\omega^{(1)}+s)g_x^{(1)}-\omega^{(2)}+\dot{g}_y^{(1)}+\frac{g_y^{(1)}}{2}\frac{d\omega}{dt}\left(\frac{1}{\omega-m}-\frac{1}{\omega}\right),$$
$$(\omega+m)(2f_x^{(2)}+(f_x^{(1)})^2+(f_y^{(1)})^2)+(\omega-m)(2g_x^{(2)}+(g_x^{(1)})^2+(g_y^{(1)})^2)=0\,, \quad (3.142)$$

which has as solutions

$$f_x^{(2)} = \frac{m^2\ddot{a}}{8a\omega^4}-\frac{m\ddot{a}}{8a\omega^3}-\frac{5m^4\dot{a}^2}{16a^2\omega^6}+\frac{5m^3\dot{a}^2}{16a^2\omega^5}+\frac{3m^2\dot{a}^2}{32a^2\omega^4}-\frac{m\dot{a}^2}{8a^2\omega^3}+\frac{5m^2s^2}{8\omega^4}-\frac{ms^2}{2\omega^3}-\frac{s^2}{8\omega^2}\,,$$
$$\omega^{(2)} = \frac{-m^2s^2}{2\omega^3}+\frac{s^2}{2\omega}+\frac{5m^4\dot{a}^2}{8a^2\omega^5}-\frac{3m^2\dot{a}^2}{8a^2\omega^3}-\frac{m^2\ddot{a}}{4a\omega^3}\,, \quad (3.143)$$

and $g_x^{(2)}(m,s)=f_x^{(2)}(-m,-s)$. On the other hand, taking the imaginary part of the equations, we have

$$(\omega-m)(g_y^{(2)}-f_y^{(2)}) = (\omega^{(1)}-s)f_y^{(1)}+\dot{f}_x^{(1)}+\frac{f_x^{(1)}}{2}\frac{d\omega}{dt}\left(\frac{1}{\omega+m}-\frac{1}{\omega}\right),$$
$$(\omega+m)(g_y^{(2)}-f_y^{(2)}) = -(\omega^{(1)}+s)g_y^{(1)}-\dot{g}_x^{(1)}-\frac{g_x^{(1)}}{2}\frac{d\omega}{dt}\left(\frac{1}{\omega-m}-\frac{1}{\omega}\right). \quad (3.144)$$

As before, this system contains an arbitrariness in its solution,

$$f_y^{(2)}=B+\frac{5m^2s\dot{a}}{4a\omega^4}-\frac{s\dot{a}}{2a\omega^2}-\frac{\dot{s}}{2\omega^2}\,,\qquad g_y^{(2)}=B\,, \quad (3.145)$$

where now B is a linear combination of second order adiabatic terms. By imposing again the condition $F^{(2)}(m,s)=G^{(2)}(-m,-s)$, one finds

$$f_y^{(2)}=\frac{5m^2s\dot{a}}{8a\omega^4}-\frac{s\dot{a}}{4a\omega^2}-\frac{\dot{s}}{4\omega^2}\,, \quad (3.146)$$

and $g_y^{(2)}(m,s)=f_y^{(2)}(-m,-s)$.

Third and fourth adiabatic order

The same procedure can be repeated for all orders. The real part of the expansion is totally determined by the system of equations (3.133), while every imaginary part contains and arbitrariness that can be solved by fixing the condition $F^{(n)}(m,s) = G^{(n)}(-m,-s)$. The third and fourth order terms of the expansion are explicitly written in Appendix D

3.7 Renormalization of the stress-energy tensor $\langle T_{\mu\nu}\rangle$ and the bilinear $\langle\bar{\psi}\psi\rangle$

Recall from previous sections that the classical stress-energy tensor in a FLRW spacetime (3.37) has two independent components. For a Dirac field, they are (no sum on i),

$$T_0^0 = \frac{i}{2}\left(\bar{\psi}\gamma^0\frac{\partial\psi}{\partial t} - \frac{\partial\bar{\psi}}{\partial t}\gamma^0\psi\right) , \quad T_i^i = \frac{i}{2a}\left(\bar{\psi}\gamma^i\frac{\partial\psi}{\partial x^i} - \frac{\partial\bar{\psi}}{\partial x^i}\gamma^i\psi\right) . \tag{3.147}$$

We define the vacuum state $|0\rangle$ as $B_{\vec{k},\lambda}|0\rangle \equiv D_{\vec{k},\lambda}|0\rangle \equiv 0$, and denote any expectation value on this vacuum as e.g. $\langle T_{\mu\nu}\rangle \equiv \langle 0|T_{\mu\nu}|0\rangle$. In the quantum theory, the vacuum expectation values of the stress-energy tensor take the form (recall (3.41), (3.42) and (3.58), (3.59))

$$\langle T_{00}\rangle = \frac{1}{2\pi^2a^3}\int_0^\infty dkk^2\rho_k(t) , \quad \rho_k(t) \equiv 2i\left(h_k^I\frac{\partial h_k^{I*}}{\partial t} + h_k^{II}\frac{\partial h_k^{II*}}{\partial t}\right) , \tag{3.148}$$

and

$$\langle T_{ii}\rangle = \frac{1}{2\pi^2a}\int_0^\infty dkk^2p_k(t) , \quad p_k(t) \equiv -\frac{2k}{3a}(h_k^Ih_k^{II*} + h_k^{I*}h_k^{II}) . \tag{3.149}$$

The above formal expressions contain quartic, quadratic, and logarithmic UV divergences, which turn out to be independent of the particular quantum state. These divergences are similar to those described in [181]. To characterize them, one plugs in (3.148)-(3.149) the adiabatic expansion of h_k^I and h_k^{II}, given in equation (3.131). We shall see that, in the presence of a Yukawa interaction, all adiabatic orders up to the fourth one generate UV divergences. This is different to what happens in the case of a free field, where the divergences only appear at zeroth and second adiabatic orders [68]. In general, adiabatic renormalization proceeds by subtracting those adiabatic terms from the integrand of the expectation values, producing a formal finite quantity. There are two important considerations regarding these subtractions. First, they must refer to all contributions of a given adiabatic term of fixed (adiabatic) order, othe rwise general covariance is not maintained. And second, one subtracts only the minimum number of terms required to get a finite result [157].

We now proceed to calculate the renormalized expressions for the energy density and pressure.

Renormalized energy density

We start by performing the adiabatic expansion of the energy density in momentum space (3.148)

$$\rho_k = \rho_k^{(0)} + \rho_k^{(1)} + \rho_k^{(2)} + \rho_k^{(3)} + \rho_k^{(4)} + \dots \ , \tag{3.150}$$

where $\rho_k^{(n)}$ is of nth adiabatic order. The adiabatic terms producing UV divergences (after integration in momenta) are

$$\rho_k^{(0)} = -2\omega \ , \tag{3.151}$$

$$\rho_k^{(1)} = -\frac{2ms}{\omega} \ , \tag{3.152}$$

$$\rho_k^{(2)} = -\frac{\dot a^2 m^4}{4a^2\omega^5} + \frac{\dot a^2 m^2}{4a^2\omega^3} + \frac{m^2 s^2}{\omega^3} - \frac{s^2}{\omega} \ , \tag{3.153}$$

$$\rho_k^{(3)} = \frac{5\dot a^2 m^5 s}{4a^2\omega^7} - \frac{7\dot a^2 m^3 s}{4a^2\omega^5} + \frac{\dot a^2 ms}{2a^2\omega^3} - \frac{\dot a m^3 \dot s}{2a\omega^5} + \frac{\dot a m \dot s}{2a\omega^3} - \frac{m^3 s^3}{\omega^5} + \frac{ms^3}{\omega^3} \ , \tag{3.154}$$

$$\begin{aligned}\rho_k^{(4)} = {} & \frac{105\dot a^4 m^8}{64a^4\omega^{11}} - \frac{91\dot a^4 m^6}{32a^4\omega^9} + \frac{81\dot a^4 m^4}{64a^4\omega^7} - \frac{\dot a^4 m^2}{16a^4\omega^5} - \frac{7\dot a^2 m^6 \ddot a}{8a^3\omega^9} + \frac{5\dot a^2 m^4 \ddot a}{4a^3\omega^7} - \frac{3\dot a^2 m^2 \ddot a}{8a^3\omega^5} \\ & - \frac{35\dot a^2 m^6 s^2}{8a^2\omega^9} + \frac{15\dot a^2 m^4 s^2}{2a^2\omega^7} - \frac{m^4\ddot a^2}{16a^2\omega^7} - \frac{27\dot a^2 m^2 s^2}{8a^2\omega^5} + \frac{m^2\ddot a^2}{16a^2\omega^5} + \frac{\dot a^2 s^2}{4a^2\omega^3} + \frac{\dot a m^4 a^{(3)}}{8a^2\omega^7} \\ & - \frac{\dot a m^2 a^{(3)}}{8a^2\omega^5} + \frac{5\dot a m^4 s\dot s}{2a\omega^7} - \frac{3\dot a m^2 s\dot s}{a\omega^5} + \frac{\dot a s\dot s}{2a\omega^3} + \frac{5m^4 s^4}{4\omega^7} - \frac{3m^2 s^4}{2\omega^5} - \frac{m^2\dot s^2}{4\omega^5} + \frac{s^4}{4\omega^3} + \frac{\dot s^2}{4\omega^3} \ ,\end{aligned} \tag{3.155}$$

where we have used the notation $a^{(3)} \equiv d^3a/dt^3, a^{(4)} \equiv d^4a/dt^4$, etc.

We note that if we turn off the Yukawa coupling, we recover the results obtained in previous sections [68]. The Yukawa interaction produces new contributions and, in particular, we have now non-zero terms at first and third adiabatic orders. The physical meaning of them will be given later on. Note here that in the UV limit, $\rho_k^{(0)} \sim k$, $(\rho_k^{(1)} + \rho_k^{(2)}) \sim k^{-1}$, and $(\rho_k^{(3)} + \rho_k^{(4)}) \sim k^{-3}$. This indicates that subtracting the zeroth-order term will cancel the natural quartic divergence of the stress-energy tensor, subtracting up to second order will cancel also the quadratic divergence, and subtracting up to fourth order will cancel the logarithmic divergence. Therefore, defining the adiabatic subtraction terms as

$$\langle T_{00}\rangle_{Ad} \equiv \frac{1}{2\pi^2 a^3}\int_0^\infty dk k^2 (\rho_k^{(0)} + \rho_k^{(1)} + \rho_k^{(2)} + \rho_k^{(3)} + \rho_k^{(4)}) \equiv \frac{1}{2\pi^2 a^3}\int_0^\infty dk k^2 \rho_k^{(0-4)} \ , \tag{3.156}$$

the renormalized 00-component of the stress-energy tensor is

$$\langle T_{00}\rangle_{ren} \equiv \langle T_{00}\rangle - \langle T_{00}\rangle_{Ad} = \frac{1}{2\pi^2 a^3}\int_0^\infty dk k^2(\rho_k - \rho_k^{(0-4)}) \,. \tag{3.157}$$

This integral is, by construction, finite.

Renormalized pressure

The method proceeds in the same way for the pressure. The renormalized ii-component of the stress-energy tensor is given by

$$\langle T_{ii}\rangle_{ren} \equiv \langle T_{ii}\rangle - \langle T_{ii}\rangle_{Ad} = \frac{1}{2\pi^2 a}\int_0^\infty dk k^2(p_k - p_k^{(0-4)}) \,, \tag{3.158}$$

where $p_k^{(0-4)} \equiv p_k^{(0)} + p_k^{(1)} + p_k^{(2)} + p_k^{(3)} + p_k^{(4)}$, and

$$\langle T_{ii}\rangle_{Ad} \equiv \frac{1}{2\pi^2 a}\int_0^\infty dk k^2 p_k^{(0-4)} \,. \tag{3.159}$$

The corresponding adiabatic terms for the pressure are

$$p_k^{(0)} = -\frac{2\omega}{3} + \frac{2m^2}{3\omega}, \tag{3.160}$$

$$p_k^{(1)} = \frac{2ms}{3\omega} - \frac{2m^3 s}{3\omega^3}, \tag{3.161}$$

$$p_k^{(2)} = -\frac{5\dot a^2 m^6}{12a^2\omega^7} + \frac{\dot a^2 m^4}{2a^2\omega^5} - \frac{\dot a^2 m^2}{12a^2\omega^3} + \frac{m^4\ddot a}{6a\omega^5} - \frac{m^2\ddot a}{6a\omega^3} + \frac{m^4 s^2}{\omega^5} - \frac{4m^2 s^2}{3\omega^3} + \frac{s^2}{3\omega}, \tag{3.162}$$

$$\begin{aligned} p_k^{(3)} = {} & -\frac{35\dot a^2 m^7 s}{12a^2\omega^9} - \frac{5\dot a^2 m^5 s}{a^2\omega^7} + \frac{9\dot a^2 m^3 s}{4a^2\omega^5} - \frac{\dot a^2 ms}{6a^2\omega^3} - \frac{5m^5 s\ddot a}{6a\omega^7} - \frac{5\dot a m^5\dot s}{6a\omega^7} + \frac{7m^3 s\ddot a}{6a\omega^5} + \frac{7\dot a m^3\dot s}{6a\omega^5} \\ & -\frac{ms\ddot a}{3a\omega^3} - \frac{\dot a m\dot s}{3a\omega^3} - \frac{5m^5 s^3}{3\omega^7} + \frac{8m^3 s^3}{3\omega^5} + \frac{m^3\ddot s}{6\omega^5} - \frac{ms^3}{\omega^3} - \frac{m\ddot s}{6\omega^3}\,, \end{aligned} \tag{3.163}$$

$$\begin{aligned} p_k^{(4)} = {} & \frac{385\dot a^4 m^{10}}{64a^4\omega^{13}} - \frac{791\dot a^4 m^8}{64a^4\omega^{11}} + \frac{1477\dot a^4 m^6}{192a^4\omega^9} - \frac{m^4 a^{(4)}}{24a\omega^7} - \frac{263\dot a^4 m^4}{192a^4\omega^7} + \frac{m^2 a^{(4)}}{24a\omega^5} + \frac{\dot a^4 m^2}{48a^4\omega^5} \\ & -\frac{77\dot a^2 m^8\ddot a}{16a^3\omega^{11}} + \frac{203\dot a^2 m^6\ddot a}{24a^3\omega^9} - \frac{191\dot a^2 m^4\ddot a}{48a^3\omega^7} + \frac{\dot a^2 m^2\ddot a}{3a^3\omega^5} - \frac{105\dot a^2 m^8 s^2}{8a^2\omega^{11}} + \frac{665\dot a^2 m^6 s^2}{24a^2\omega^9} + \frac{7m^6\ddot a^2}{16a^2\omega^9} \\ & -\frac{145\dot a^2 m^4 s^2}{8a^2\omega^7} - \frac{5m^4\ddot a^2}{8a^2\omega^7} + \frac{29\dot a^2 m^2 s^2}{8a^2\omega^5} + \frac{3m^2\ddot a^2}{16a^2\omega^5} - \frac{\dot a^2 s^2}{12a^2\omega^3} + \frac{7\dot a m^6 a^{(3)}}{12a^2\omega^9} - \frac{5\dot a m^4 a^{(3)}}{6a^2\omega^7} \\ & +\frac{\dot a m^2 a^{(3)}}{4a^2\omega^5} + \frac{35m^6 s^2\ddot a}{12a\omega^9} + \frac{35\dot a m^6 s\dot s}{6a\omega^9} - \frac{5m^4 s^2\ddot a}{a\omega^7} - \frac{10\dot a m^4 s\dot s}{a\omega^7} + \frac{9m^2 s^2\ddot a}{4a\omega^5} + \frac{9\dot a m^2 s\dot s}{2a\omega^5} \\ & -\frac{s^2\ddot a}{6a\omega^3} - \frac{\dot a s\dot s}{3a\omega^3} + \frac{35m^6 s^4}{12\omega^9} - \frac{65m^4 s^4}{12\omega^7} - \frac{5m^4 s\ddot s}{6\omega^7} - \frac{5m^4\dot s^2}{12\omega^7} + \frac{11m^2 s^4}{4\omega^5} + \frac{m^2 s\ddot s}{\omega^5} + \frac{m^2\dot s^2}{3\omega^5} \\ & -\frac{s^4}{4\omega^3} - \frac{s\ddot s}{6\omega^3} + \frac{\dot s^2}{12\omega^3}\,. \end{aligned} \tag{3.164}$$

As before, we see that in the UV limit, $p_k^{(0)} \sim k$, $(p_k^{(1)} + p_k^{(2)}) \sim k^{-1}$, and $(p_k^{(3)} + p_k^{(4)}) \sim k^{-3}$. Subtracting the zeroth-order term eliminates the quartic divergence, subtracting up to second order removes the quadratic divergence, and subtracting up to fourth order removes the logarithmic divergence. If the Yukawa interaction is removed, we recover again the results of previous sections [68]. The interaction produces also non-zero contributions to the first and third adiabatic orders.

Renormalization of $\langle\bar{\psi}\psi\rangle$

As argued in section 3.5, we are also interested in computing the renormalized expectation value $\langle\bar{\psi}\psi\rangle_{ren}$. The formal (unrenormalized) expression for this quantity is

$$\langle\bar{\psi}\psi\rangle = \frac{-1}{\pi^2 a^3}\int_0^\infty dk k^2 \langle\bar{\psi}\psi\rangle_k \ , \quad \langle\bar{\psi}\psi\rangle_k \equiv |h_k^I|^2 - |h_k^{II}|^2 \ . \tag{3.165}$$

We define the corresponding terms in the adiabatic expansion as $\langle\bar{\psi}\psi\rangle_k = \langle\bar{\psi}\psi\rangle_k^{(0)} + \langle\bar{\psi}\psi\rangle_k^{(1)} + \langle\bar{\psi}\psi\rangle_k^{(2)} + \langle\bar{\psi}\psi\rangle_k^{(3)} + \ldots$. Due to the Yukawa interaction, ultraviolet divergences arrive until the third adiabatic order. In general, we have

$$\langle\bar{\psi}\psi\rangle_k^{(n)} = \frac{\omega+m}{2\omega}\left(|F|^2\right)^{(n)} - \frac{\omega-m}{2\omega}\left(|G|^2\right)^{(n)} \ . \tag{3.166}$$

From here, we obtain

$$\langle\bar{\psi}\psi\rangle_k^{(0)} = \frac{m}{\omega} \ , \tag{3.167}$$

$$\langle\bar{\psi}\psi\rangle_k^{(1)} = \frac{s}{\omega} - \frac{m^2 s}{\omega^3} \ , \tag{3.168}$$

$$\langle\bar{\psi}\psi\rangle_k^{(2)} = -\frac{5\dot{a}^2 m^5}{8a^2\omega^7} + \frac{7\dot{a}^2 m^3}{8a^2\omega^5} - \frac{\dot{a}^2 m}{4a^2\omega^3} + \frac{m^3\ddot{a}}{4a\omega^5} - \frac{m\ddot{a}}{4a\omega^3} + \frac{3m^3 s^2}{2\omega^5} - \frac{3ms^2}{2\omega^3} \ , \tag{3.169}$$

$$\begin{aligned}\langle\bar{\psi}\psi\rangle_k^{(3)} = {} & \frac{35\dot{a}^2 m^6 s}{8a^2\omega^9} - \frac{15\dot{a}^2 m^4 s}{2a^2\omega^7} + \frac{27\dot{a}^2 m^2 s}{8a^2\omega^5} - \frac{\dot{a}^2 s}{4a^2\omega^3} - \frac{5m^4 s\ddot{a}}{4a\omega^7} - \frac{5\dot{a}m^4\dot{s}}{4a\omega^7} + \frac{3m^2 s\ddot{a}}{2a\omega^5} + \frac{2\dot{a}m^2\dot{s}}{a\omega^5} \\ & - \frac{s\ddot{a}}{4a\omega^3} - \frac{3\dot{a}\dot{s}}{4a\omega^3} - \frac{5m^4 s^3}{2\omega^7} + \frac{3m^2 s^3}{\omega^5} + \frac{m^2\ddot{s}}{4\omega^5} - \frac{s^3}{2\omega^3} - \frac{\ddot{s}}{4\omega^3} \ . \end{aligned} \tag{3.170}$$

The adiabatic prescription leads then to

$$\langle\bar{\psi}\psi\rangle_{ren} = \langle\bar{\psi}\psi\rangle - \langle\bar{\psi}\psi\rangle_{Ad} = \frac{-1}{\pi^2 a^3}\int_0^\infty dk k^2 (\langle\bar{\psi}\psi\rangle_k - \langle\bar{\psi}\psi\rangle_k^{(0-3)}) \ . \tag{3.171}$$

In this case, we observe that in the UV limit, $(\langle\bar{\psi}\psi\rangle_k^{(0)} + \langle\bar{\psi}\psi\rangle_k^{(1)}) \sim k^{-1}$, and $(\langle\bar{\psi}\psi\rangle_k^{(2)} + \langle\bar{\psi}\psi\rangle_k^{(3)}) \sim k^{-3})$. Subtracting up to first order eliminates the quadratic divergence, and up to third order removes the logarithmic one.

Our results can be generically implemented together with numerical methods to compute the renormalized expectation values $\langle T_{\mu\nu}\rangle_{ren}$ and $\langle\bar{\psi}\psi\rangle_{ren}$. On the other hand, we would like to briefly comment that a higher-order adiabatic expansion also serves to generate asymptotic analytical expressions for the renormalized stress-energy tensor in some special situations. This happens in spacetime regions where the relevant modes always evolve adiabatically. For instance, if we approximate the form of the exact modes $\{h_k^I, h_k^{II}\}$ by their higher-order adiabatic expansion, we can find in a very straightforward way an analytic approximation for the renormalized quantities in the adiabatic regime, as in the example given in Appendix E. Outside the adiabatic regime one should use numerical methods to find the exact modes and plug them in the generic renormalized expressions obtained above.

3.8 UV divergences and renormalization counterterms in the effective action

The ultraviolet divergent terms of the adiabatic subtractions can be univocally related to particular counterterms in the Lagrangian density including the background gravity-scalar sector. By writing

$$\begin{aligned}\mathscr{L} &= \mathscr{L}_m + \sqrt{-g}\left[\frac{1}{2}g^{\mu\nu}\nabla_\mu\Phi\nabla_\nu\Phi - \sum_{i=1}^{4}\frac{\lambda_i}{i!}\Phi^i - \xi_1 R\Phi - \frac{1}{2}\xi_2 R\Phi^2 - \frac{1}{8\pi G}\Lambda + \frac{1}{16\pi G}R\right] \\ &+ \sqrt{-g}\left[\frac{1}{2}\delta Z g^{\mu\nu}\nabla_\mu\Phi\nabla_\nu\Phi - \sum_{i=1}^{4}\frac{\delta\lambda_i}{i!}\Phi^i - \delta\xi_1 R\Phi - \frac{1}{2}\delta\xi_2 R\Phi^2 - \frac{1}{8\pi}\delta\Lambda + \frac{1}{16\pi}\delta G^{-1}R\right] ,\end{aligned} \tag{3.172}$$

the equations of motion for the scalar field read

$$\begin{aligned}&(1+\delta Z)\Box\Phi + (\lambda_1+\delta\lambda_1) + (\lambda_2+\delta\lambda_2)\Phi + (\lambda_3+\delta\lambda_3)\frac{1}{2}\Phi^2 \\ &+ \frac{1}{3!}(\lambda_4+\delta\lambda_4)\Phi^3 + (\xi_1+\delta\xi_1)R + (\xi_2+\delta\xi_2)R\Phi = -g_Y\langle\bar{\psi}\psi\rangle .\end{aligned} \tag{3.173}$$

From (3.171), we can write the identity

$$\langle\bar{\psi}\psi\rangle = \langle\bar{\psi}\psi\rangle_{ren} + \frac{1}{\pi^2 a^3}\int_0^\infty dk k^2 (\langle\bar{\psi}\psi\rangle_k^{(0)} + \langle\bar{\psi}\psi\rangle_k^{(1)} + \langle\bar{\psi}\psi\rangle_k^{(2)} + \langle\bar{\psi}\psi\rangle_k^{(3)}) , \tag{3.174}$$

where $\langle\bar{\psi}\psi\rangle_{ren}$ is finite and the remaining integrals at the right-hand-side of (3.174) are the adiabatic subtraction terms. As we shall see, the ultraviolet divergences of the adiabatic subtraction terms can be removed by counterterms of the form: $\delta Z\Box\Phi$, $\delta\lambda_1$, $\delta\lambda_2\Phi$, $\delta\lambda_3\Phi^2$, $\delta\lambda_4\Phi^3$, $\delta\xi_1 R$, and $\delta\xi_2 R\Phi$. To deal with the UV-divergent subtraction terms we use dimensional regularization [53]. We can check that the (covariantly) regulated divergences take the same form as the above covariant counterterms. For $\langle\bar{\psi}\psi\rangle^{(0)}$ we have (n denotes the space-time dimension)

$$\langle\bar{\psi}\psi\rangle^{(0)} = -\frac{1}{\pi^2 a^3}\int_0^\infty dk k^2 \left(-\frac{m}{\omega(t)}\right) \to -\frac{1}{\pi^2 a^3}\int_0^\infty dk k^{n-2}\left(-\frac{m}{\omega(t)}\right) = \frac{m^3}{2\pi^2(n-4)} + ... \tag{3.175}$$

where we will retain only the poles at $n=4$. This divergence can be absorbed by $\delta\lambda_1$. Additionally, we also have

$$\langle\bar{\psi}\psi\rangle^{(1)} = -\frac{1}{\pi^2 a^3}\int_0^\infty dk k^{n-2}\left(-\frac{s(t)k^2}{\omega^3(t)a(t)^2}\right) = \frac{3g_Y m^2}{2\pi^2(n-4)}\Phi(t) + \cdots\ . \tag{3.176}$$

This divergence of adiabatic order one can be absorbed by $\delta\lambda_2$. The divergences of adiabatic order two

$$\langle\bar{\psi}\psi\rangle^{(2)} = -\frac{m}{24\pi^2(n-4)}R + \frac{3mg_Y^2}{2\pi^2(n-4)}\Phi^2(t) + \cdots \tag{3.177}$$

can also be eliminated by $\delta\xi_1$ and $\delta\lambda_3$. Finally, the three divergences of adiabatic order three

$$\langle\bar{\psi}\psi\rangle^{(3)} = \frac{g_Y}{4\pi^2(n-4)}\Box\Phi(t) + \frac{g_Y}{24\pi^2(n-4)}R\Phi(t) + \frac{g_Y^3}{2\pi^2(n-4)}\Phi^3(t) + \cdots \tag{3.178}$$

are absorbed by δZ, $\delta\xi_2$ and $\delta\lambda_4$.

On the other hand, the tensorial equations are

$$\begin{aligned}&\frac{1}{8\pi}\left(\frac{1}{G}+\delta G^{-1}\right)G^{\mu\nu} + \frac{1}{8\pi}\left(\frac{\Lambda}{G}+\delta\Lambda\right)g^{\mu\nu} + (1+\delta Z)(\nabla^\mu\Phi\nabla^\nu\Phi - \frac{1}{2}g^{\mu\nu}\nabla^\rho\Phi\nabla_\rho\Phi)\\ + \ &g^{\mu\nu}\sum_{i=1}^{4}\frac{(\lambda_i+\delta\lambda_i)}{i!}\Phi^i - 2\sum_{i=1}^{2}\frac{\xi_i+\delta\xi_i}{i!}(G^{\mu\nu}\Phi^i - g^{\mu\nu}\Box\Phi^i + \nabla^\mu\nabla^\nu\Phi^i) = -\langle T_m^{\mu\nu}\rangle\ ,\end{aligned} \tag{3.179}$$

and we find similar cancelations. However, two extra divergences appear. Focusing, for simplicity, at zeroth adiabatic order, we have

$$\begin{aligned}-\frac{1}{2\pi^2 a^3}\int_0^\infty dk k^{n-2}\rho_k^{(0)} &\approx \frac{m^4}{8\pi^2}\frac{1}{n-4}\ ,\\ -\frac{1}{2\pi^2 a^3}\int_0^\infty dk k^{n-2}a^2 p_k^{(0)} &\approx -\frac{m^4 a^2}{8\pi^2}\frac{1}{n-4}\ .\end{aligned} \tag{3.180}$$

At first adiabatic order we encounter the following divergences

$$\begin{aligned}-\frac{1}{2\pi^2 a^3}\int_0^\infty dk k^{n-2}\rho_k^{(1)} &\approx \frac{m^3 s}{2\pi^2}\frac{1}{n-4}\ ,\\ -\frac{1}{2\pi^2 a^3}\int_0^\infty dk k^{n-2}a^2 p_k^{(1)} &\approx -\frac{m^3 a^2 s}{2\pi^2}\frac{1}{n-4}\ .\end{aligned} \tag{3.181}$$

At second adiabatic order we find these divergences

$$
\begin{aligned}
-\frac{1}{2\pi^2a^3}\int_0^\infty dkk^{n-2}\rho_k^{(2)} &\approx \frac{m^2}{8\pi^2}\frac{1}{n-4}\frac{\dot{a}^2}{a^2}+\frac{3m^2}{4\pi^2}\frac{1}{n-4}s^2 \, , \\
-\frac{1}{2\pi^2a^3}\int_0^\infty dkk^{n-2}a^2p_k^{(2)} &\approx -\frac{m^2a^2}{24\pi^2}\frac{1}{n-4}\left(2\frac{\ddot{a}}{a}+\frac{\dot{a}^2}{a^2}\right)-\frac{3a^2m^2s^2}{4\pi^2}\frac{1}{n-4} \, . \qquad (3.182)
\end{aligned}
$$

At third adiabatic order we get the following divergences ($H\equiv \dot{a}/a$)

$$
\begin{aligned}
-\frac{1}{2\pi^2a^3}\int_0^\infty dkk^{n-2}\rho_k^{(3)} &\approx \frac{m}{12\pi^2}\frac{1}{n-4}\left[3H^2s+3H\dot{s}+6s^3\right] \, , \qquad (3.183) \\
-\frac{1}{2\pi^2a^3}\int_0^\infty dkk^{n-2}a^2p_k^{(3)} &\approx -\frac{ma^2}{12\pi^2}\frac{1}{n-4}\left[\ddot{s}+2H\dot{s}+\left(H^2+2\frac{\ddot{a}}{a}\right)s+6s^3\right] \, .
\end{aligned}
$$

Finally, at fourth adiabatic order the divergences are

$$
\begin{aligned}
-\frac{1}{2\pi^2a^3}\int_0^\infty dkk^{n-2}\rho_k^{(4)} &\approx \frac{1}{8\pi^2}\frac{1}{n-4}\left[H^2s^2+s^4+2H\dot{s}s+\dot{s}^2\right] \, , \qquad (3.184) \\
-\frac{1}{2\pi^2a^3}\int_0^\infty dkk^{n-2}a^2p_k^{(4)} &\approx -\frac{a^2}{8\pi^2}\frac{1}{n-4}\left[s^4+\left(H^2+2\frac{\ddot{a}}{a}\right)\frac{s^2}{3}-\frac{\dot{s}^2}{3}+\frac{4}{3}Hs\dot{s}+\frac{2}{3}s\ddot{s}\right] \, .
\end{aligned}
$$

All the above divergent expressions arising from the Yukawa interaction can be written covariantly as

$$
\begin{aligned}
\langle T_{\mu\nu}\rangle_{Ad}^{(0)} &\approx \frac{m^4}{8\pi^2(n-4)}g_{\mu\nu} \, , \qquad (3.185) \\
\langle T_{\mu\nu}\rangle_{Ad}^{(1)} &\approx \frac{g_Y\Phi m^3}{2\pi^2(n-4)}g_{\mu\nu} \, , \qquad (3.186) \\
\langle T_{\mu\nu}\rangle_{Ad}^{(2)} &\approx \frac{3g_Y^2\Phi^2m^2}{4\pi^2(n-4)}g_{\mu\nu}-\frac{m^2}{24\pi^2(n-4)}G_{\mu\nu} \, , \qquad (3.187) \\
\langle T_{\mu\nu}\rangle_{Ad}^{(3)} &\approx -\frac{mg_Y}{12\pi^2(n-4)}\left[G_{\mu\nu}\Phi-\Box\Phi g_{\mu\nu}+\nabla_\mu\nabla_\nu\Phi-6g_Y^2\Phi^3g_{\mu\nu}\right] \, , \qquad (3.188) \\
\langle T_{\mu\nu}\rangle_{Ad}^{(4)} &\approx \frac{-g_Y^2}{24\pi^2(n-4)}\left[G_{\mu\nu}\Phi^2-g_{\mu\nu}\Box\Phi^2+\nabla_\mu\nabla_\nu\Phi^2-6(\nabla_\mu\Phi\nabla_\nu\Phi-\frac{1}{2}g_{\mu\nu}\nabla_\rho\Phi\nabla^\rho\Phi)-3g_Y^2\Phi^4g_{\mu\nu}\right] \qquad (3.189)
\end{aligned}
$$

and can be consistently removed (including also the divergences for $\langle\bar{\psi}\psi\rangle$) by the renormalization parameters

$$
\delta\Lambda=-\frac{m^4}{\pi(n-4)} \, , \quad \delta G^{-1}=\frac{m^2}{3\pi(n-4)} \, , \quad \delta Z=-\frac{g_Y^2}{4\pi^2(n-4)} \, , \qquad (3.190)
$$

$$
\delta\lambda_1=-\frac{m^3g_Y}{2\pi^2(n-4)} \, , \quad \delta\lambda_2=-\frac{3m^2g_Y^2}{2\pi^2(n-4)} \, , \quad \delta\lambda_3=-\frac{3mg_Y^3}{\pi^2(n-4)} \, , \quad \delta\lambda_4=-\frac{3g_Y^4}{\pi^2(n-4)} \, , \qquad (3.191)
$$

$$
\delta\xi_1=-\frac{mg_Y}{24\pi^2(n-4)} \, , \quad \delta\xi_2=-\frac{g_Y^2}{24\pi^2(n-4)} \, . \qquad (3.192)
$$

We remark that the set of needed counterterms are all possible counterterms having couplings with non-negative mass dimension, up to Newton's coupling constant. This is also in agreement with the results in perturbative QFT in flat spacetime. The renormalizability of the Yukawa interaction $g_Y \varphi\bar{\psi}\psi$ of a quantized massive scalar field φ with a massive quantized Dirac field ψ requires to add terms of the form $\frac{\lambda Z_\lambda}{4!}\varphi^4$, $\frac{\kappa Z_\kappa}{3!}\varphi^3$, and also a term linear in φ [178]. The presence of a curved background would require to add the terms $\xi_1 R\varphi$ and $\xi_2 R\varphi^2$. We note that a term of the form $\xi_2 R\varphi^2$ is required by renormalization for a purely quantized scalar field φ if a self-interaction term of the form $\frac{\lambda}{4!}\varphi^4$ appears in the bare Lagrangian density [52, 54]. Here we have found that the Yukawa interaction demands the presence of the renormalized terms $\xi_1 R\varphi$ and $\xi_2 R\varphi^2$ (as well as the terms $\lambda_i \varphi^i$), even if they are not present in the bare Lagrangian density. Similar counterterms have been identified in the approach in Ref. [37].

Therefore, the tentative semiclassical equations presented in Section 3.5 should be reconsidered to include the above required counterterms. In terms of the renormalized parameters we have

$$\begin{aligned}&\frac{1}{8\pi G}(G^{\mu\nu}+\Lambda g^{\mu\nu})+(\nabla^\mu\Phi\nabla^\nu\Phi-\frac{1}{2}g^{\mu\nu}\nabla^\rho\Phi\nabla_\rho\Phi+V(\Phi)g^{\mu\nu})\\ - \;& 2\sum_{i=1}^{2}\frac{\xi_i}{i!}(G^{\mu\nu}\Phi^i-g^{\mu\nu}\Box\Phi^i+\nabla^\mu\nabla^\nu\Phi^i)=-\langle T_m^{\mu\nu}\rangle_{ren}\ ,\end{aligned} \tag{3.193}$$

and

$$\Box\Phi+\frac{\partial V}{\partial\Phi}+\xi_1 R+\xi_2 R\Phi=-g_Y\langle\bar{\psi}\psi\rangle_{ren}\ , \tag{3.194}$$

where the potential $V(\Phi)$ should contain the terms

$$V(\Phi)=\lambda_1\Phi+\frac{\lambda_2}{2}\Phi^2+\frac{\lambda_3}{3!}\Phi^3+\frac{\lambda_4}{4!}\Phi^4\ . \tag{3.195}$$

Obviously, additional terms, not required by renormalization, can be added to the potential if one adopts an effective field theory viewpoint. Some of the renormalized parameters (Λ, ξ_1, λ_1, $\cdots$) could take, by fine tuning, zero values. We do not consider these issues in this work.

3.9 Conformal anomaly

In this section we will analyze the massless limit of the theory and work out the conformal anomaly. In the massless limit the classical action of the theory enjoys invariance under the conformal transformations

$$g_{\mu\nu}(x)\to\Omega^2(x)g_{\mu\nu}(x)\ ,\quad \Phi(x)\to\Omega^{-1}(x)\Phi(x)\ , \tag{3.196}$$

with

$$\psi(x) \to \Omega^{-3/2}(x)\psi(x)\ , \qquad \bar{\psi}(x) \to \Omega^{-3/2}(x)\bar{\psi}(x). \tag{3.197}$$

Variation of the action yields the identity

$$g^{\mu\nu}T^m_{\mu\nu} + \Phi\frac{1}{\sqrt{-g}}\frac{\delta S_m}{\delta\Phi} = 0\ , \tag{3.198}$$

which, in our case, turns out to be $g^{\mu\nu}T_{\mu\nu} - g_Y\Phi\bar{\psi}\psi = 0$. At the quantum level the theory will lose its conformal invariance as a consequence of renormalization [which respects general covariance and hence (3.116)] and generates an anomaly

$$g^{\mu\nu}\langle T^m_{\mu\nu}\rangle_{ren} - g_Y\Phi\langle\bar{\psi}\psi\rangle_{ren} = C_f \neq 0\ . \tag{3.199}$$

C_f is independent of the quantum state and depends only on local quantities of the external fields.

To calculate the conformal anomaly in the adiabatic regularization method, we have to start with a massive field and take the massless limit at the end of the calculation. Therefore,

$$C_f = g^{\mu\nu}\langle T^m_{\mu\nu}\rangle_{ren} - g_Y\Phi\langle\bar{\psi}\psi\rangle_{ren} = \lim_{m\to 0} m(\langle\bar{\psi}\psi\rangle_{ren} - \langle\bar{\psi}\psi\rangle^{(4)})\ . \tag{3.200}$$

Since the divergences of the stress-energy tensor have terms of fourth adiabatic order, the adiabatic subtractions for $\langle\bar{\psi}\psi\rangle$ should also include them. The fourth order subtraction term, which produces a non-zero finite contribution when $m \to 0$, is codified in $\langle\bar{\psi}\psi\rangle^{(4)}$. The term $m\langle\bar{\psi}\psi\rangle_{ren}$ vanishes when $m \to 0$. The remaining piece produces the anomaly [recall (3.165)-(3.166)]

$$\begin{aligned} C_f &= -\lim_{m\to 0}\frac{m}{\pi^2 a^3}\int_0^\infty dk k^2\Big(-\frac{(\omega+m)}{2\omega}[F^{(4)} + F^{(4)*} + F^{(1)}F^{(3)*} + F^{(1)*}F^{(3)} + |F^{(2)}|^2] \\ &+ \frac{(\omega-m)}{2\omega}[G^{(4)} + G^{(4)*} + G^{(1)}G^{(3)*} + G^{(1)*}G^{(3)} + |G^{(2)}|^2]\Big)\ . \end{aligned} \tag{3.201}$$

Applying the adiabatic expansion computed in Section 3.6 and doing the integrals we obtain

$$C_f = \frac{a^{(4)}}{80\pi^2 a} + \frac{s^2\ddot{a}}{8\pi^2 a} + \frac{\ddot{a}^2}{80\pi^2 a} + \frac{3s\dot{s}\dot{a}}{4\pi^2 a} + \frac{s^2\dot{a}^2}{8\pi^2 a^2} + \frac{3\dot{a}a^{(3)}}{80\pi^2 a^2} - \frac{\dot{a}^2\ddot{a}}{60\pi^2 a^3} + \frac{s\ddot{s}}{4\pi^2} + \frac{\dot{s}^2}{8\pi^2} + \frac{s^4}{8\pi^2}\ . \tag{3.202}$$

Since C_f is an scalar, we must be able to rewrite the above result as a linear combination of covariant scalar terms made out of the metric, the Riemann tensor, covariant derivatives, and the external scalar field Φ. Our result is

$$C_f = \frac{1}{2880\pi^2}\left[-11\left(R_{\alpha\beta}R^{\alpha\beta} - \frac{1}{3}R^2\right) + 6\Box R\right] + \frac{g_Y^2}{8\pi^2}\left[\nabla^\mu\Phi\nabla_\mu\Phi + 2\Phi\Box\Phi + \frac{1}{6}\Phi^2 R + g_Y^2\Phi^4\right]. \tag{3.203}$$

The same result is obtained by using the results of Section 3.7. C_f can be re-expressed as [recall (3.157)-(3.158)]

$$C_f = \lim_{m\to 0} \frac{-1}{2\pi^2 a^3} \int_0^\infty dk k^2 \left(\rho_k^{(0-4)} - 3p_k^{(0-4)} - 2(s(t)+m)\langle\bar{\psi}\psi\rangle_k^{(0-3)} \right) . \qquad (3.204)$$

Performing the integrals we get exactly (3.202) and hence (3.203).

In Appendix F we have computed the conformal anomaly for a massless scalar field ϕ with conformal coupling to the scalar curvature $\xi = 1/6$, and with a Yukawa-type interaction of the form $g_Y^2 \Phi^2 \phi^2$. Adiabatic regularization predicts the following conformal anomaly

$$C_s = \frac{1}{2880\pi^2} \left[\Box R - \left(R^{\mu\nu} R_{\mu\nu} - \frac{1}{3} R^2 \right) \right] - \frac{h^2}{48\pi^2} (\Phi \Box \Phi + \nabla^\mu \Phi \nabla_\mu \Phi + \frac{3h^2}{2} \Phi^4) . \qquad (3.205)$$

In the absence of Yukawa interaction ($h = 0$, $g_Y = 0$) we reproduce the well-known trace anomaly for both scalar and spin-1/2 fields (restricted to our FLRW spacetime) [48]. We recall that the trace anomaly is generically given for a conformal free field of spin 0, 1/2 or 1 in terms of three coefficients

$$g^{\mu\nu} \langle T_{\mu\nu} \rangle_{ren} = a C_{\mu\nu\rho\sigma} C^{\mu\nu\rho\sigma} + bG + c \Box R \ , \qquad (3.206)$$

where $C_{\mu\nu\rho\sigma}$ is the Weyl tensor and $G = R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} - 4 R_{\mu\nu} R^{\mu\nu} + R^2$ is proportional to the Euler density. The coefficients a and b are independent of the renormalization scheme and are given by [76, 114]

$$\begin{aligned} a &= \frac{1}{120(4\pi)^2} (N_s + 6N_f + 12N_v) \ , \\ b &= \frac{-1}{360(4\pi)^2} (N_s + 11N_f + 62N_v) \ , \end{aligned} \qquad (3.207)$$

where N_s is the number of real scalar fields, N_f is the number of Dirac fields, and N_v is the number of vector fields. Our results with $g_Y = 0$ fit the values in (3.207). [We note that in the FLRW spacetime of adiabatic regularization the Weyl tensor vanishes identically]. In contrast, the coefficient c depends in general on the particular renormalization scheme [183]. A local counterterm proportional to R^2 in the action can modify the coefficient c. For instance, for vector fields the point-splitting and the dimensional regularization method predict different values for c.

When the Yukawa interaction is added, the general form of the conformal anomaly is

$$\begin{aligned} g^{\mu\nu} \left\langle T^m_{\mu\nu} \right\rangle_{ren} + \Phi \frac{1}{\sqrt{-g}} \langle \frac{\delta S_m}{\delta \Phi} \rangle_{ren} &= a C_{\mu\nu\rho\sigma} C^{\mu\nu\rho\sigma} + bG + c \Box R \\ &+ d g_Y^2 \nabla^\mu \Phi \nabla_\mu \Phi + e g_Y^2 \Phi \Box \Phi + f g_Y^2 \Phi^2 R + g g_Y^4 \Phi^4 \ . \end{aligned} \qquad (3.208)$$

Now, the coefficients f and g are independent of the renormalization scheme but d and e are not. The finite Lagrangian counterterms required by the renormalizability of the Yukawa interaction obtained in previous sections,

$$\frac{\delta Z}{2} g^{\mu\nu}\nabla_\mu\Phi\nabla_\nu\Phi - \frac{\delta\xi_2}{2}R\Phi^2 - \frac{\delta\lambda_4}{4!}\Phi^4 \ , \tag{3.209}$$

might alter the values of the coefficients d and e, but not the coefficients f and g. Note that, due to classical conformal invariance, one should consider only those counterterms having dimensionless coupling parameters. Therefore, our results for the f and g coefficients are

$$f = \frac{1}{3(4\pi)^2}N_f \ , \quad g = \frac{-1}{3(4\pi)^2}\left(\frac{3}{2}N_s - 6N_f\right) \ . \tag{3.210}$$

Finally, to show explicitly that the above coefficients are independent on the renormalization scheme, we will compute them using the heat-kernel method given in [157], by means of the one-loop effective action.

Consistency with the heat-kernel results

The conformal anomaly for a field $\phi^j(x)$ obeying the second-order wave equation

$$\left[\delta^i_j g^{\mu\nu}\nabla_\mu\nabla_\nu + Q^i_j(x)\right]\phi^j = 0 \ , \tag{3.211}$$

is given by

$$C = \pm\frac{1}{(4\pi)^2} tr\ E_2(x) \ , \tag{3.212}$$

where $E_2(x)$ is the second Seeley-DeWitt coefficient. The minus sign is for bosons and the plus sign is for fermions. These coefficients are local, scalar functions of $Q(x)$ and the curvature tensor. E_2 is given by

$$E_2 = \left(-\frac{1}{30}\Box R + \frac{1}{72}R^2 - \frac{1}{180}R^{\mu\nu}R_{\mu\nu} + \frac{1}{180}R^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma}\right)I + \frac{1}{12}W^{\mu\nu}W_{\mu\nu} + \frac{1}{2}Q^2 - \frac{1}{6}RQ + \frac{1}{6}\Box Q \ , \tag{3.213}$$

where $W_{\mu\nu} = [\nabla_\mu, \nabla_\nu]$. For a single massless scalar field with $\xi = 1/6$ and an interaction of the form $h^2\phi^2\Phi^2$ we have

$$Q = \frac{1}{6}R + h^2\Phi^2 \ , \tag{3.214}$$

and $W_{\mu\nu} = 0$. For a spatially flat FLRW universe we get

$$C = \frac{1}{2880\pi^2}\left[\Box R - \left(R^{\mu\nu}R_{\mu\nu} - \frac{1}{3}R^2\right)\right] - \frac{h^2}{48\pi^2}(\Phi\Box\Phi + \nabla^\mu\Phi\nabla_\mu\Phi + \frac{3h^2}{2}\Phi^4) \ , \tag{3.215}$$

in full agreement with the result (3.205) obtained using adiabatic regularization.

For a single massless Dirac field with a Yukawa interaction we have $[(i\gamma^\mu\nabla_\mu - g_Y\Phi)\psi = 0]$

$$Q = \left(\frac{1}{4}R + g_Y^2\Phi^2\right)I + ig_Y\gamma^\mu\nabla_\mu\Phi \ , \tag{3.216}$$

and

$$W_{\mu\nu} = -iR_{\mu\nu}{}^{\alpha\beta}\Sigma_{\alpha\beta} = -\frac{1}{8}R_{\mu\nu}{}^{\alpha\beta}[\gamma_\alpha,\gamma_\beta] \ . \tag{3.217}$$

Using the properties of the trace of products of gamma matrices, we get

$$C = \frac{1}{2880\pi^2}\left[-11\left(R_{\alpha\beta}R^{\alpha\beta} - \frac{1}{3}R^2\right) + 6\Box R\right] + \frac{g_Y^2}{8\pi^2}\left[-\frac{1}{3}\nabla^\mu\Phi\nabla_\mu\Phi + \frac{2}{3}\Phi\Box\Phi + \frac{1}{6}\Phi^2 R + g_Y^2\Phi^4\right] . \tag{3.218}$$

The above result reproduces the coefficients f and g obtained from adiabatic regularization. We note that there is a mismatch in the coefficients d and e. These are, however, the coefficients that might depend on the renormalization scheme.

3.10 Adiabatic vs DeWitt-Schwinger formalisms. Spin-0 fields

We study now the equivalence between the adiabatic formalism presented above and the DeWitt-Schinger one. As we have already mentioned in the introduction, the equivalence between both methods has been checked in [25, 47] for scalar fields by direct computation. We present here an alternative and simpler approach for scalar fields that will allow us to prove the equivalence for spin-1/2 fields.

Adiabatic regularization

The general wave equation for a scalar field ϕ in a curved space-time is $(\Box + m^2 + \xi R)\phi = 0$, where m is the mass of the field and ξ is the coupling of the field to the scalar curvature R. If the field propagates in a FLRW space-time [for simplicity we shall assume a spatially flat universe with metric $ds^2 = dt^2 - a^2(t)d\vec{x}^2$], it can be naturally expanded in the form

$$\phi(x) = \int d^3k \left[A_{\vec{k}} f_{\vec{k}}(\vec{x},t) + A^\dagger_{\vec{k}} f^*_{\vec{k}}(\vec{x},t)\right] \ , \tag{3.219}$$

where the field modes $f_{\vec{k}}$ are

$$f_{\vec{k}}(t,\vec{x}) = \frac{e^{i\vec{k}\cdot\vec{x}}}{\sqrt{2(2\pi)^3 a^3(t)}} h_k(t) \ . \tag{3.220}$$

These modes are assumed to obey the normalization condition with respect to the conserved Klein-Gordon product. This condition translates into a Wronskian-type condition for the modes: $h_k^* \dot{h}_k - \dot{h}_k^* h_k = -2i$, where the dot means derivative with respect to proper time t. Adiabatic renormalization is based on a generalized WKB-type asymptotic expansion of the modes according to the ansatz

$$h_k(t) \sim \frac{1}{\sqrt{W_k(t)}} e^{-i\int^t W_k(t')dt'} , \tag{3.221}$$

which solves the Wronskian condition. One then expands W_k in an adiabatic series, in which each contribution is determined by the number of time derivatives of the expansion factor $a(t)$

$$W_k(t) = \omega^{(0)}(t) + \omega^{(2)}(t) + \omega^{(4)}(t) + \dots , \tag{3.222}$$

where the leading term $\omega^{(0)}(t) \equiv \omega(t) = \sqrt{k^2/a^2(t) + m^2}$ is the usual physical frequency. Higher order contributions can be univocally obtained by iteration (for details, see Appendix G), which come from introducing (3.221) into the equation of motion for the modes. As we have learnt from previus sections, this adiabatic expansion (3.222) is basic to identify and remove the UV divergences of the expectation values of the stress-energy tensor.

The adiabatic expansion of the modes can be easily translated to an expansion of the two-point function $\langle\phi(x)\phi(x')\rangle \equiv G(x,x')$ at coincidence $x = x'$:

$$G_{Ad}(x,x) = \frac{1}{2(2\pi)^3 a^3} \int d^3\vec{k}\, [\omega^{-1} + (W^{-1})^{(2)} + (W^{-1})^{(4)} + \dots] . \tag{3.223}$$

As remarked above the expansion must be truncated to the minimal adiabatic order necessary to cancel all UV divergences that appear in the formal expression of the vacuum expectation value that one wishes to compute. The calculation of the renormalized variance $\langle\phi^2\rangle$ requires only second adiabatic order, given by

$$(W^{-1})^{(2)} = \frac{m^2\dot{a}^2}{2a^2\omega^5} + \frac{m^2\ddot{a}}{4a\omega^5} - \frac{5m^4\dot{a}^2}{8a^2\omega^7} + \frac{3(\frac{1}{6}-\xi)(\dot{a}^2 + a\ddot{a})}{a^2\omega^3} . \tag{3.224}$$

The renormalization of the vacuum expectation value of the stress-energy tensor needs up to fourth adiabatic order subtraction. The corresponding fourth order contribution $(W^{-1})^{(4)}$ has 30 terms and can be found in [152]. Therefore, the asymptotic two point function at coincident points, truncated to fourth adiabatic order, can be rewritten as

$${}^{(4)}G_{Ad}(x,x) = \frac{1}{2(2\pi)^3 a^3} \int d^3\vec{k} \left[\frac{1}{\omega} + \frac{(\frac{1}{6}-\xi)R}{2\omega^3} + \frac{m^2\dot{a}^2}{2a^2\omega^5} + \frac{m^2\ddot{a}}{4a\omega^5} - \frac{5m^4\dot{a}^2}{8a^2\omega^7} + (W^{-1})^{(4)}\right] \tag{3.225}$$

where we have taken into account that $R = 6[\dot{a}^2/a^2 + \ddot{a}/a]$ in FLRW universes.

We note that only the first two terms in (3.225) are divergent. The remaining terms can be integrated exactly in momenta producing well-defined finite geometric quantities. Taking into account that $\omega = (\vec{k}^2/a^2 + m^2)^{1/2}$, the integration of the second order adiabatic terms is independent of the mass and gives

$$\frac{1}{2(2\pi)^3 a^3}\int d^3\vec{k}\left[\frac{m^2\dot{a}^2}{2a^2\omega^5} + \frac{m^2\ddot{a}}{4a\omega^5} - \frac{5m^4\dot{a}^2}{8a^2\omega^7}\right] = \frac{R}{288\pi^2}\ . \tag{3.226}$$

The integration of the fourth order terms turns out to be also a well-defined geometrical quantity

$$\frac{1}{2(2\pi)^3 a^3}\int d^3\vec{k}\,(W^{-1})^{(4)} = \frac{a_2}{16\pi^2 m^2}\ , \tag{3.227}$$

where

$$a_2 = \frac{1}{2}\left[\xi - \frac{1}{6}\right]^2 R^2 - \frac{1}{6}\left[\frac{1}{5} - \xi\right]\Box R - \frac{1}{180}(R_{\mu\nu}R^{\mu\nu} - R_{\mu\nu\gamma\delta}R^{\mu\nu\gamma\delta})\ , \tag{3.228}$$

is just the coincident point limit $a_2(x) \equiv \lim_{x\to x'} a_2(x,x')$ of the second DeWitt coefficient $a_2(x,x')$ [71]. (We note that, for our conformally flat space-times, the Weyl tensor vanishes and $R_{\mu\nu\gamma\delta}R^{\mu\nu\gamma\delta} = 2R_{\mu\nu}R^{\mu\nu} - \frac{1}{3}R^2$).

In summary, the asymptotic two-point function for a scalar field at coincidence truncated at fourth adiabatic order is given by

$${}^{(4)}G_{Ad}(x,x) = \frac{1}{4\pi^2 a^3}\int_0^\infty dk k^2\left[\frac{1}{\omega} + \frac{(\frac{1}{6}-\xi)R}{2\omega^3}\right] + \frac{R}{288\pi^2} + \frac{a_2}{16\pi^2 m^2}\ , \tag{3.229}$$

where the formal divergent term can be understood, for future purposes, as the point-splitting limit

$$\frac{1}{4\pi^2 a^3}\int_0^\infty dk k^2\left[\frac{1}{\omega} + \frac{(\frac{1}{6}-\xi)R}{2\omega^3}\right] \equiv \lim_{|\Delta\vec{x}|\to 0}\frac{1}{4\pi^2 a^3}\int_0^\infty dk k^2\frac{\sin(k|\Delta\vec{x}|)}{k|\Delta\vec{x}|}\left[\frac{1}{\omega} + \frac{(\frac{1}{6}-\xi)R}{2\omega^3}\right]\ . \tag{3.230}$$

Local momentum-space representation and DeWitt-Schwinger expansion

An alternative asymptotic expansion of the two-point function in momentum space was introduced by Bunch and Parker in [54]. It was proposed to aim at extending to curved space the standard momentum-space methods of perturbation theory for interacting fields in Minkowski space. This way the standard Minkowskian propagator of a scalar free field in momentum space $(-k^2+m^2)^{-1}$ is replaced by a series expansion. The Fourier transform leading to local-momentum space is crucially performed with respect to Riemann normal coordinates y^μ around a given point x', which constitutes the best possible approximation in curved space to the inertial coordinates of Minkowski space. In contrast to adiabatic regularization, the method is valid for an arbitrary space-time. It does not serve to (adiabatically) expand the mode functions, which are otherwise highly ambiguous in a general background. The method works directly with the two-point functions, which are regarded as the basic buildings blocks of the renormalization process.

The covariant expansion of the two-point function $G_{DS}(x,x')$, obeying the equation

$$(\Box_x + m^2 + \xi R)G_{DS}(x,x') = -|g(x)|^{-1/2}\delta(x-x') \, , \tag{3.231}$$

is defined in the local-momentum space

$$G_{DS}(x,x') = \frac{-i|g(x)|^{-1/4}}{(2\pi)^4}\int d^4k \; e^{iky}\bar{G}(k) \, , \tag{3.232}$$

where $ky \equiv k_0 y^0 - \vec{k}\vec{y}$ (note that $y^\mu(x')=0$), by the series

$$\begin{aligned} \bar{G}(k) &= \frac{1}{-k^2+m^2} + \frac{(\frac{1}{6}-\xi)R}{(-k^2+m^2)^2} + \frac{i(\frac{1}{6}-\xi)}{2}R_{;\alpha}\frac{\partial}{\partial k_\alpha}\frac{1}{(-k^2+m^2)^2} + \frac{1}{3}a_{\alpha\beta}\frac{\partial}{\partial k_\alpha}\frac{\partial}{\partial k_\beta}(-k^2+m^2)^{-2} \\ &+ \left[\left(\frac{1}{6}-\xi\right)^2 R^2 + \frac{2}{3}a^\alpha_\alpha\right]\frac{1}{(-k^2+m^2)^3} + \dots \, , \end{aligned} \tag{3.233}$$

where

$$a_{\alpha\beta} = \frac{(\xi-\frac{1}{6})}{2}R_{;\alpha\beta} + \frac{1}{120}R_{;\alpha\beta} - \frac{1}{40}\Box R_{\alpha\beta} - \frac{1}{30}R_\alpha{}^\lambda R_{\lambda\beta} + \frac{1}{60}R_{\kappa\alpha\lambda\beta}R^{\kappa\lambda} + \frac{1}{60}R^{\lambda\mu\kappa}{}_\alpha R_{\lambda\mu\kappa\beta} \, . \tag{3.234}$$

To compare this local-momentum expansion with the adiabatic one introduced in the previous subsection we have to convert the momentum-space four-dimensional integrals into three-dimensional integrals. After performing the k^0 integration in the complex plane, where the poles in $\bar{G}(k)$ at $k_0 = \pm\sqrt{\vec{k}^2+m^2}$ have been displaced in the same way as the analogous Green function in Minkowski space-time, one gets tridimensional integrals. Since all Green functions have the same UV divergences we can perform the contour k^0 integration using, for instance, the Feynman prescription for displacing the poles. The result is, up to fourth adiabatic order,

$$
\begin{aligned}
{}^{(4)}G_{DS}(x,x') &= \frac{|g(x)|^{-1/4}}{2(2\pi)^3}\int d^3k\, e^{-i(\vec{k}\vec{y}-\sqrt{\vec{k}^2+m^2}\,y^0)} \qquad (3.235)\\
&\times \left[\frac{a_0}{(\vec{k}^2+m^2)^{1/2}}+\frac{a_1(x,x')(1-iy^0\omega)}{2(\vec{k}^2+m^2)^{3/2}}+\frac{3a_2(x,x')(1-iy^0\omega-(y^0)^2\omega^2/3)}{4(\vec{k}^2+m^2)^{5/2}}\right]\\
&= \frac{|g(x)|^{-1/4}}{(2\pi)^2|\vec{y}|}\int_0^\infty dk\,k\sin(k|\vec{y}|)e^{i\sqrt{\vec{k}^2+m^2}\,y^0} \qquad (3.236)\\
&\times \left[\frac{a_0}{(\vec{k}^2+m^2)^{1/2}}+\frac{a_1(x,x')(1-iy^0\omega)}{2(\vec{k}^2+m^2)^{3/2}}+\frac{3a_2(x,x')(1-iy^0\omega-(y^0)^2\omega^2/3)}{4(\vec{k}^2+m^2)^{5/2}}\right],
\end{aligned}
$$

where $a_0(x,x') \equiv 1$ and, to fourth adiabatic order,

$$
\begin{aligned}
a_1(x,x') &= \left[\frac{1}{6}-\xi\right]R(x')+\frac{1}{2}\left[\frac{1}{6}-\xi\right]R_{;\alpha}(x')y^\alpha-\frac{1}{3}a_{\alpha\beta}(x')y^\alpha y^\beta\,,\\
a_2(x,x') &= \frac{1}{2}\left[\frac{1}{6}-\xi\right]^2R^2(x')+\frac{1}{3}a^\alpha_\alpha(x')\,, \qquad (3.237)
\end{aligned}
$$

which turn out to be the first DeWitt coefficients. The integrals can be worked out analytically and (3.235) gives the first three terms in the DeWitt-Schwinger expansion of the two-point function [48]

$$
\begin{aligned}
{}^{(4)}G_{DS}(x,x') &= \frac{|g(x)|^{-1/4}}{4\pi^2}\left[\frac{m}{\sqrt{-2\sigma}}K_1(m\sqrt{-2\sigma})+\frac{a_1(x,x')}{2}K_0(m\sqrt{-2\sigma})\right.\\
&\left.+\frac{a_2(x,x')}{4m}\sqrt{-2\sigma}K_1(m\sqrt{-2\sigma})\right], \qquad (3.238)
\end{aligned}
$$

where $\sigma(x,x')$ is half the square of the geodesic distance between x and x', i.e., $\sigma(x,x') = \frac{1}{2}y_\mu y^\mu = ((y^0)^2-\vec{y}^2)/2$, and K are the modified Bessel functions of second kind.

It is also important to note that the factor $|g(x)|^{-1/4}$ in the above expressions is evaluated in Riemann normal coordinates with origin at x'. The biscalar that reduces to $|g(x)|^{-1/4}$ in arbitrary coordinates is $\Delta^{1/2}(x,x')$, where $\Delta(x,x')$ is the Van Vleck - Morette determinant, defined as

$$\Delta(x,x') = -|g(x)|^{-1/2}\det[-\partial_\mu\partial_{\nu'}\sigma(x,x')]|g(x')|^{-1/2}\,. \qquad (3.239)$$

These expressions fit identically with the conventional definition of the DeWitt-Schwinger expansion, as first stressed in [54], which is usually written as

$$G_{DS}(x,x') \equiv \frac{\Delta^{1/2}(x,x')}{16\pi^2}\int_0^\infty \frac{ids}{(is)^2}\exp\left(-im^2s+\frac{\sigma}{2is}\right)F(x,x';is)\,, \qquad (3.240)$$

with

$$F(x,x';is) = a_0 + a_1(x,x')is + a_2(x,x')(is)^2 + \ldots\,, \qquad (3.241)$$

where $a_0 = 1, a_1, a_2, ...$ are the DeWitt coefficients. To sum up, the Bunch-Parker local momentum-space expansion turns out to be the momentum-space version of the DeWitt-Schwinger expansion of the two-point function.

Comparison between ${}^{(4)}G_{Ad}(x,x)$ and ${}^{(4)}G_{DS}(x,x)$

To compare the expression (3.238) for $G_{DS}(x,x')$ with the result of adiabatic regularization (3.229) we have to take the coincident limit $x = x'$ and restrict our analysis to a spatially flat FLRW universe $ds^2 = dt^2 - a^2(t)d\vec{x}^2$. The comparison is not trivial since in the DeWitt-Schwinger formalism the point-splitting is studied in terms of the geodesic distance σ. As a first approximation, the normal Riemann coordinates in our FLRW space-time are $\vec{y} \approx a\Delta\vec{x}$. To rigorously compare with the adiabatic expansion we need the higher order relations between the physical coordinates $(t,\vec{x})$ and the normal Riemann coordinates $(y^0,\vec{y})$. The following relations (with $H = \dot{a}/a$) hold [51]

$$y^0 = \Delta t + \frac{1}{2}a^2\Delta\vec{x}^2 H + \frac{1}{3}a^2\Delta\vec{x}^2\Delta t\left(\frac{R}{12} + H^2\right) + \dots\,, \tag{3.242}$$

$$y^i = a\Delta x^i\left[1 + H\Delta t + \frac{1}{6}a^2\Delta\vec{x}^2H^2 + \frac{\Delta t^2}{3}\left(\frac{R}{6} - H^2\right) + \dots\right]. \tag{3.243}$$

Moreover,

$$-2\sigma = -\Delta t^2 + a^2\Delta\vec{x}^2 + a^2\Delta\vec{x}^2 H\Delta t + \frac{1}{3}a^2\Delta\vec{x}^2\Delta t^2\left(\frac{R}{6} - H^2\right) + \frac{a^4\Delta\vec{x}^4}{12}H^2 + \dots\,, \tag{3.244}$$

where, in order to compare to our previous result using the adiabatic regularization, we can just take $\Delta t = 0$ without loss of generality and retain the point splitting in $\Delta\vec{x}$.

A useful identity for our purposes, using (3.244) at temporal coincidence $\Delta t = 0$, is

$$\frac{1}{-2\sigma} = \frac{1}{a^2\Delta\vec{x}^2} - \frac{H^2}{12} + O(\Delta\vec{x}^2)\,. \tag{3.245}$$

Note also that the factor $|g(x)|^{-1/4}$ in (3.238) is evaluated in Riemann normal coordinates with origin at x' so we can expand $|g(x,x')|^{-1/4} = \Delta^{1/2}(x,x') = 1 - \frac{1}{12}R_{\mu\nu}y^\mu y^\nu + \dots$. Another useful relation can be derived using this last result with formulas (3.242) - (3.243) [note also that $R_{00} = 3\frac{\ddot{a}}{a}; R_{ii} = -a^2\left(\frac{\ddot{a}}{a} + 2H^2\right)$],

$$|g(x)|^{-1/4} = 1 - \left[2H^2 + \frac{\ddot{a}}{a}\right]\frac{\sigma}{6} + O(\sigma^{3/2})\,. \tag{3.246}$$

Taking into account (3.245) and (3.246), the zeroth order contribution to ${}^{(4)}G_{DS}(x,x)$ can be reexpressed as

$$\lim_{x\to x'} \frac{|g(x)|^{-1/4} m}{(2\pi)^2\sqrt{-2\sigma}} K_1(m\sqrt{-2\sigma}) = \lim_{x\to x'} |g(x)|^{-1/4}\left[-\frac{1}{8\pi^2\sigma} + O(\log(-\sigma))\right] \tag{3.247}$$

$$= \frac{R}{288\pi^2} + \lim_{\Delta\vec{x}\to 0} \frac{m}{4\pi^2 a|\Delta\vec{x}|} K_1(ma|\Delta\vec{x}|) \tag{3.248}$$

$$= \frac{R}{288\pi^2} + \lim_{\Delta\vec{x}\to 0} \frac{1}{4\pi^2 a^3}\int_0^\infty dk k^2 \frac{\sin(k|\Delta\vec{x}|)}{k|\Delta\vec{x}|}\frac{1}{\omega} \tag{3.249}$$

Furthermore, the second order contribution is

$$\lim_{x\to x'} \frac{|g(x)|^{-1/4}}{4\pi^2}\frac{a_1(x,x')}{2} K_0(m\sqrt{-2\sigma}) = \lim_{x\to x'} |g(x)|^{-1/4} \times O(\log(-\sigma)) \tag{3.250}$$

$$= \lim_{\Delta\vec{x}\to 0} \frac{1}{4\pi^2}\frac{\left(\frac{1}{6}-\xi\right)R}{2} K_0(ma|\Delta\vec{x}|) \tag{3.251}$$

$$= \lim_{\Delta\vec{x}\to 0} \frac{1}{4\pi^2 a^3}\int_0^\infty dk k^2 \frac{\sin(k|\Delta\vec{x}|)}{k|\Delta\vec{x}|}\frac{\left(\frac{1}{6}-\xi\right)R}{2\omega^3} \tag{3.252}$$

while the fourth adiabatic term is given by

$$\lim_{x\to x'} \frac{|g(x)|^{-1/4}}{4\pi^2}\frac{a_2(x,x')}{4m}\sqrt{-2\sigma} K_1(m\sqrt{-2\sigma}) = \frac{a_2(x)}{16\pi^2 m^2} \ . \tag{3.253}$$

To sum up, we finally get

$${}^{(4)}G_{DS}(x,x) = \lim_{|\Delta\vec{x}|\to 0} \frac{1}{4\pi^2 a^3}\int_0^\infty dk k^2 \frac{\sin(k|\Delta\vec{x}|)}{k|\Delta\vec{x}|}\left[\frac{1}{\omega} + \frac{(\frac{1}{6}-\xi)R}{2\omega^3}\right] + \frac{R}{288\pi^2} + \frac{a_2(x)}{16\pi^2 m^2} \ . \tag{3.254}$$

By direct comparison with (3.229) and (3.230) we obtain

$${}^{(4)}G_{Ad}(x,x) = {}^{(4)}G_{DS}(x,x) \ . \tag{3.255}$$

Equivalence for $\langle T_{\mu\nu}\rangle$

For the sake of simplicity it is now convenient to restrict ourselves to the case $\xi = 1/6$. The reason for which we focus on this particular case is because the spin-1/2 case turns out to be completely analogous, so that it is an illustrative example. In this situation the trace of the stress-energy tensor can be expressed as $\langle T\rangle = m^2\langle\phi^2\rangle$. The equivalence $\langle T\rangle^{Ad} = \langle T\rangle^{DS}$, and hence $\langle T_{\mu\nu}\rangle^{Ad} = \langle T_{\mu\nu}\rangle^{DS}$ (i.e., $c_3 = 0$, according to the definitions and arguments given in Sec. 5.1), comes directly from the equivalence ${}^{(4)}G_{Ad}(x,x) = {}^{(4)}G_{DS}(x,x)$, since

$$\langle T\rangle^{Ad} - \langle T\rangle^{DS} = m^2\left[{}^{(4)}G_{DS}(x,x) - {}^{(4)}G_{Ad}(x,x)\right] = 0 \ . \tag{3.256}$$

For a general ξ, one can compute the stress-energy tensor by acting on the symmetric part of $G(x,x')-{}^{(4)}G(x,x')$ with a certain nonlocal operator, $\langle T_{\mu\nu}(x)\rangle = \lim_{x'\to x} D_{\mu\nu}(x,x')[G(x,x')-{}^{(4)}G(x,x')]$ [48, 59, 71]. In Section 3.13 we will show the equivalence ${}^{(4)}G_{Ad}(x,x') = {}^{(4)}G_{DS}(x,x')$, which immediately implies $\langle T_{\mu\nu}\rangle^{Ad} = \langle T_{\mu\nu}\rangle^{DS}$ for a general ξ.

3.11 Adiabatic vs DeWitt-Schwinger formalisms. Spin-1/2 fields

We extend the ideas presented in the previous section for the spin 1/2 case.

Adiabatic regularization

The first step in the adiabatic regularization is to define an asymptotic expansion of the field modes. The expansion can be regarded as definitions of approximate particle states in an expanding universe in the limit of infinitely slow expansion.

Spin-1/2 fields obey the Dirac equation (3.5). Taking a spatially flat FLRW spacetime, with line element $ds^2 = dt^2 - a^2(t)d\vec{x}^2$, it yields (3.6). Recall that we work with the Dirac-Pauli representation of the Minkowskian Dirac matrices (3.7). For a given comoving momentum $\vec{k}$, the basic independent (normalized) spinor solutions are (3.8) and (3.13). where $k \equiv |\vec{k}|$ and ξ_λ are constant and normalized two-component spinor $\xi_\lambda^\dagger \xi_{\lambda'} = \delta_{\lambda'\lambda}$. See section 3.2 for more details. In this decomposition, h_k^I and h_k^{II} are two particular time-dependent functions obeying the following coupled differential equations (3.9), and we take, as in previous sections, the following self-consistent expansion for the field modes

$$h_k^I(t) \sim \sqrt{\frac{\omega+m}{2\omega}}\, e^{-i\int^{t'}\Omega(t')dt'}F(t) \quad , \quad h_k^{II}(t) \sim \sqrt{\frac{\omega-m}{2\omega}}\, e^{-i\int^{t'}\Omega(t')dt'}G(t)\,, \tag{3.257}$$

where $\omega \equiv \omega^0 \equiv \sqrt{(k/a(t))^2+m^2}$ is the frequency of the mode and the time-dependent functions $\Omega(t)$, $F(t)$ and $G(t)$ are expanded adiabatically as

$$\Omega(t) = \sum_{n=0}^{\infty} \omega^{(n)}(t)\,, \quad F(t) = \sum_{n=0}^{\infty} F^{(n)}(t)\,, \quad G(t) = \sum_{n=0}^{\infty} G^{(n)}(t)\,. \tag{3.258}$$

$\omega^{(n)}$, $F^{(n)}$ and $G^{(n)}$ are functions of adiabatic order n, which means that they contain n derivatives of the scale factor $a(t)$. We impose $F^{(0)} = G^{(0)} \equiv 1$ at zeroth order to recover the Minkowskian solutions for $a(t) = 1$. We can solve $\omega^{(n)}$, $F^{(n)}$ and $G^{(n)}$ for $n > 1$ by direct substitution of the ansatz (3.257) into (3.9) and solving the system of equations order by order. We also have to impose, as an additional order by order requirement, the normalization condition $|h_k^I(t)|^2 + |h_k^{II}(t)|^2 = 1$. For details, see section 3.2, from (3.19) to (3.23). The adiabatic series obtained in this way contain ambiguities. The ambiguities disappear in the adiabatic expansion of physical vacuum expectation values. It is very convenient, for the sake of simplicity, to impose at all adiabatic orders the additional condition $ImG^{(n)}(m) = -ImF^{(n)}(m)$. It implies that $F^{(n)}(-m) = G^{(n)}(m)$ and removes all the ambiguities. Explicit expressions for the series expansion up to fourth adiabatic order are displayed in (3.20), (3.21), (3.22), (3.23), and Appendix A. The algorithm to obtain systematically $\omega^{(n)}$, $F^{(n)}$ and $G^{(n)}$ for any nth adiabatic order is shown in Appendix H.

In parallel with the scalar field, the adiabatic expansion of the spin-1/2 field modes can be translated to an expansion of the two-point function $\langle\psi_\alpha(x)\bar{\psi}_\beta(x')\rangle \equiv S_{\alpha\beta}(x,x')$ at coincidence $x = x'$. Moreover, since we are mainly interested in studying the stress-energy tensor we will restrict our analysis to the trace of the two-point function $\langle\bar{\psi}(x')\psi(x)\rangle = trS(x,x')$. Evaluating this at coincidence, the adiabatic expansion up to fourth order is

$$tr^{(4)}S_{Ad}(x,x) = \frac{-2}{(2\pi)^3a^3}\int d^3k\,[|g_k^{I(4)}|^2 - |g_k^{II(4)}|^2]\,, \tag{3.259}$$

where

$$\begin{aligned} g_k^{I(4)}(t) &\equiv \sqrt{\frac{\omega+m}{2\omega}}\sum_{n=0}^{4}F^{(n)}(t)\exp\left[-i\int^t\sum_{n=0}^{4}\omega^{(n)}(t')dt'\right]\,,\\ g_k^{II(4)}(t) &\equiv \sqrt{\frac{\omega-m}{2\omega}}\sum_{n=0}^{4}G^{(n)}(t)\exp\left[-i\int^t\sum_{n=0}^{4}\omega^{(n)}(t')dt'\right]\,. \end{aligned} \tag{3.260}$$

Taking into account that the trace of the stress-energy tensor can be expressed as $\langle T(x)\rangle = m\langle\bar{\psi}(x)\psi(x)\rangle$, it is very convenient for our purposes to rewrite (3.259) in terms of the expansion for the energy density and pressure (**??**),

$$tr^{(4)}S_{Ad}(x,x) = \frac{1}{(2\pi)^3a^3m}\int d^3k\,\sum_{i=0}^{2}[\rho_k^{(2i)} - 3p_k^{(2i)}]\,, \tag{3.261}$$

where,

$$\rho_k^{(0)} = -2\omega\,, \tag{3.262}$$

$$\rho_k^{(2)} = -\frac{m^4\dot{a}^2}{4\omega^5a^2} + \frac{m^2\dot{a}^2}{4\omega^3a^2}\,, \tag{3.263}$$

$$p_k^{(0)} = -\frac{2\omega}{3} + \frac{2m^2}{3\omega}\,, \tag{3.264}$$

$$p_k^{(2)} = -\frac{m^2\dot{a}^2}{12\omega^3a^2} - \frac{m^2\ddot{a}}{6\omega^3a} + \frac{m^4\ddot{a}}{6\omega^5a} + \frac{m^4\dot{a}^2}{2\omega^5a^2} - \frac{5m^6\dot{a}^2}{12\omega^7a^2}\,, \tag{3.265}$$

and the contribution of the fourth adiabatic order is itself finite and gives

$$\frac{1}{(2\pi)^3 a^3 m}\int d^3k\,[\rho^{(4)} - 3p^{(4)}] = \frac{trA_2}{16\pi^2 m}\ , \tag{3.266}$$

where A_2 turns out to be one of the DeWitt coefficients for spin-1/2 fields at coincidence [54, 157] (see next subsection)

$$-A_2(x) = a_2(\xi = 1/4)\mathbb{I} + \frac{1}{48}\Sigma_{[\alpha\beta]}\Sigma_{[\gamma\delta]}R^{\alpha\beta\lambda\xi}R^{\gamma\delta}{}_{\lambda\xi}\ . \tag{3.267}$$

In this equation $a_2(\xi = 1/4)$ is the DeWitt coefficient for a scalar field with curvature coupling $\xi = 1/4$, and

$$\Sigma_{[\alpha\beta]} \equiv \frac{1}{4}\left[\underline{\gamma}_\alpha\underline{\gamma}_\beta - \underline{\gamma}_\beta\underline{\gamma}_\alpha\right]\ . \tag{3.268}$$

Taking into account that

$$tr\{\Sigma_{[\alpha\beta]}\Sigma_{[\gamma\delta]}\} = g_{\alpha\delta}g_{\beta\gamma} - g_{\alpha\gamma}g_{\beta\delta}\ , \tag{3.269}$$

the term (3.266) accounts for the trace anomaly in the massless limit

$$-\frac{trA_2}{16\pi^2} = \frac{2}{2880\pi^2}\left[-\frac{11}{2}\left(R_{\mu\nu}R^{\mu\nu} - \frac{1}{3}R^2\right) + 3\Box R\right]\ . \tag{3.270}$$

Let us analyze in detail the lower orders. The zeroth order contribution is easy to handle

$$\frac{1}{(2\pi)^3 a^3 m}\int d^3k\,[\rho^{(0)} - 3p^{(0)}] = \frac{-m}{\pi^2 a^3}\int_0^\infty dk\,k^2\frac{1}{\omega}\ . \tag{3.271}$$

However, the second adiabatic order is more subtle. Using the stress-energy tensor conservation [which is equivalent as imposing the condition $\dot\rho_k^{(n)} + 3Hp_k^{(n)} = 0$, see (**??**)], and dimensional regularization, one can eventually arrive at the following expression

$$\frac{1}{(2\pi)^3 a^3 m}\int d^3k\,[\rho^{(2)} - 3p^{(2)}] = \lim_{n\to 4}\frac{-mR}{24\pi^2}\left[\frac{1}{n-4} + \frac{4}{3} - \log 2\right]\ . \tag{3.272}$$

Using now the identity

$$\frac{4m}{4\pi^2 a^3}\int_0^\infty dk\,k^2\frac{R}{24\omega^3} = \lim_{n\to 4}\frac{-mR}{24\pi^2}\left[\frac{1}{n-4} + 1 - \log 2\right]\ , \tag{3.273}$$

(3.272) can be finally expressed as

$$\frac{1}{(2\pi)^3 a^3 m}\int d^3k\,[\rho^{(2)} - 3p^{(2)}] = -4\left[\frac{mR}{288\pi^2} - \frac{m}{4\pi^2 a^3}\int_0^\infty dk\,k^2\frac{R}{24\omega^3}\right]\ . \tag{3.274}$$

Summing up we have

$$\begin{aligned} tr^{(4)}S_{Ad}(x,x) &= \frac{-m}{\pi^2 a^3}\int_0^\infty dk\,k^2\left[\frac{1}{\omega} - \frac{R}{24\omega^3}\right] - \frac{4mR}{288\pi^2} + \frac{trA_2}{16\pi^2 m} \\ &= -4m^{(2)}G_{Ad}(x,x)|_{\xi=1/4} + \frac{trA_2}{16\pi^2 m}\ . \end{aligned} \tag{3.275}$$

Local momentum-space representation and DeWitt-Schwinger expansion

Following [54, 157], one can construct an asymptotic expansion for the two-point function $\langle\psi(x)\bar{\psi}(x')\rangle \equiv S(x,x')$ as follows. Introduce the bispinor $\mathcal{G}(x,x')$ as

$$S(x,x') \equiv (i\underline{\gamma}^{\mu}\nabla_{\mu} + m)\mathcal{G}(x,x') \ . \tag{3.276}$$

This way we have, as desired,

$$\left(i\underline{\gamma}^{\mu}\nabla_{\mu} - m\right)S(x,x') = \left(\Box + m^2 + \frac{1}{4}R\right)[-\mathcal{G}(x,x')] = |g(x)|^{-1/2}\delta(x-x') \ , \tag{3.277}$$

where we used the identity, $\left(\underline{\gamma}^{\mu}\nabla_{\mu}\right)^2 = \Box + \frac{1}{4}R$ [157]. We can perform a Fourier expansion in Riemann normal coordinates around x', as in the scalar case,

$$\mathcal{G}(x,x') = \frac{-i|g(x)|^{-1/4}}{(2\pi)^4}\int d^4k \ e^{iky}\bar{\mathcal{G}}(k) \ . \tag{3.278}$$

The local-momentum expansion for spin-1/2 fields is basically that one for spin-0 fields taking $\xi = 1/4$, except for additional spinorial contributions. The detailed expansion can be looked up in [54, 157], and up to fourth adiabatic order reads

$$\begin{aligned}\bar{\mathcal{G}}(k) \ = \ & -\Bigg\{\frac{\mathbb{I}}{-k^2+m^2} - \frac{R\,\mathbb{I}}{12(-k^2+m^2)^2} - i\left[\frac{\mathbb{I}}{24}R_{;\alpha} + \frac{1}{12}\Sigma_{[\alpha\beta]}R^{\alpha\beta}{}_{\mu}{}^{\lambda}{}_{;\lambda}\right]\frac{\partial}{\partial k_{\alpha}}\frac{1}{(-k^2+m^2)^2} \\ & + \left[\frac{\mathbb{I}}{3}a_{\alpha\beta}(\xi=1/4) - \frac{1}{48}\Sigma_{[\alpha\beta]}(RR^{\alpha\beta}{}_{\mu\nu} + R^{\alpha\beta\lambda}{}_{\mu;\lambda\nu} + R^{\alpha\beta\lambda}{}_{\nu;\lambda\mu})\right. \\ & \left. + \frac{1}{96}\Sigma_{[\alpha\beta]}\Sigma_{[\gamma\delta]}(R^{\alpha\beta\lambda}{}_{\mu}R^{\gamma\delta}{}_{\lambda\nu} + R^{\alpha\beta\lambda}{}_{\nu}R^{\gamma\delta}{}_{\lambda\mu})\right]\frac{\partial}{\partial k_{\alpha}}\frac{\partial}{\partial k_{\beta}}(-k^2+m^2)^{-2} \\ & + \left[\left(\frac{R^2}{288} + \frac{1}{3}a^{\alpha}_{\alpha}(\xi=1/4)\right)\mathbb{I} + \frac{1}{48}\Sigma_{[\alpha\beta]}\Sigma_{[\gamma\delta]}R^{\alpha\beta\lambda\xi}R^{\gamma\delta}{}_{\lambda\xi}\right]\frac{2}{(-k^2+m^2)^3} + \ldots\Bigg\} \ ,\end{aligned} \tag{3.279}$$

The above expression for the spinor matrix $S(x,x')$ provides an asymptotic expansion of the two-point function $\langle\psi(x)\bar{\psi}(x')\rangle$, which also turns out to be equivalent to the DeWitt-Schwinger expansion [54]. Since we are mainly interested in $\langle\bar{\psi}(x)\psi(x)\rangle$ we take the trace of $S(x,x')$ in formulas above. Taking into account that $tr\left(\gamma^{\mu_1}\ldots\gamma^{\mu_{2k+1}}\right) = 0$, and after performing the contour k^0 integration, as in the scalar case, we obtain

$$tr\ {}^{(4)}S_{DS}(x,x') = -4m\frac{|g(x)|^{-1/4}}{2(2\pi)^3}\int d^3k e^{-i(\vec{k}\vec{y}-\sqrt{\vec{k}^2+m^2}y^0)}\left[\frac{1}{(\vec{k}^2+m^2)^{1/2}} - \frac{R(1-iy^0\omega)}{24(\vec{k}^2+m^2)^{3/2}} + \ldots\right] . \tag{3.280}$$

Restricting now the analysis to a spatially flat FLRW spacetime with metric $ds^2 = dt^2 - a^2(t)d\vec{x}^2$ and proceeding in parallel to the scalar case we get, at coincidence $x = x'$,

$$tr\ {}^{(4)}S_{DS}(x,x) = -4m\,{}^{(2)}G_{DS}(x,x)|_{\xi=1/4} + \frac{trA_2(x)}{16\pi^2 m} \ . \tag{3.281}$$

Comparison between $tr\ ^{(4)}S_{DS}(x,x)$ and $tr\ ^{(4)}S_{Ad}(x,x)$ and equivalence of $\langle T_{\mu\nu}\rangle$

It is clear from our previous results that we have a complete agreement between $tr\ ^{(4)}S_{DS}(x,x)$ and $tr\ ^{(4)}S_{Ad}(x,x)$:

$$tr\ ^{(4)}S_{DS}(x,x) = tr\ ^{(4)}S_{Ad}(x,x) = -\frac{m}{\pi^2 a^3}\int_0^\infty dk\, k^2\left[\frac{1}{\omega}-\frac{R}{24\omega^3}\right]-\frac{4mR}{288\pi^2}+\frac{trA_2(x)}{16\pi^2 m}\ . \tag{3.282}$$

As argued in section 5.1, and taking into account that $\langle T\rangle = m\langle\bar{\psi}\psi\rangle$, the equivalence $\langle T\rangle^{Ad} = \langle T\rangle^{DS}$ for spin-1/2 fields, and hence $\langle T_{\mu\nu}\rangle^{Ad} = \langle T_{\mu\nu}\rangle^{DS}$, can be simply derived from (3.282).

3.12 Extension of the equivalence to higher orders

The results obtained in previous sections suggest that the equivalence may go beyond the fourth adiabatic order, i.e., the order required to prove the equivalence of the renormalized expectation values of the stress-energy tensor. We have checked by computed assisted methods that our fundamental relations $^{(4)}G_{Ad}(x,x) = {}^{(4)}G_{DS}(x,x)$ and $tr\ ^{(4)}S_{Ad}(x,x) = tr\ ^{(4)}S_{DS}(x,x)$ are also valid at sixth adiabatic order. In the former case we have

$$^{(6)}G_{Ad}(x,x) = {}^{(6)}G_{DS}(x,x) = \frac{1}{4\pi^2 a^3}\int_0^\infty dk k^2\left[\frac{1}{\omega}+\frac{(\frac{1}{6}-\xi)R}{2\omega^3}\right]+\frac{R}{288\pi^2}+\frac{a_2}{16\pi^2 m^2}+\frac{a_3}{16\pi^2 m^4}\ , \tag{3.283}$$

where the value obtained for the purely sixth adiabatic order contribution matches exactly with the third order DeWitt coefficient a_3. The general expression for the coefficient a_3, which has 28 terms, was first obtained in [103, 173], and can also be found in [157] (see Chapter 3, Sec. 3.6). We note that the above agreement is consistent with that found in [121, 134, 135] in terms of the sixth order adiabatic approximation for the renormalized stress-energy tensor of scalar fields.

We have also tested the equivalence at sixth adiabatic order for spin-1/2 fields

$$tr\ ^{(6)}S_{DS}(x,x) = tr\ ^{(6)}S_{Ad}(x,x) = -\frac{m}{\pi^2 a^3}\int_0^\infty dk\, k^2\left[\frac{1}{\omega}-\frac{R}{24\omega^3}\right]-\frac{4mR}{288\pi^2}+\frac{trA_2}{16\pi^2 m}+\frac{trA_3}{16\pi^2 m^3}\ . \tag{3.284}$$

The adiabatic method produces the result

$$trA_3 = -\frac{2}{21}\frac{\dot{a}^4}{a^4}\frac{\ddot{a}}{a}+\frac{8}{21}\frac{\dot{a}^2}{a^2}\frac{\ddot{a}^2}{a^2}-\frac{4\ddot{a}^3}{45a^3}+\frac{2}{21}\frac{\dot{a}^3}{a^3}\frac{\dddot{a}}{a}-\frac{2}{5}\frac{\dot{a}\ddot{a}\dddot{a}}{a^3}-\frac{\dddot{a}^2}{210a^2}-\frac{\dot{a}^2\ddddot{a}}{15a^3}+\frac{\ddot{a}\ddddot{a}}{105a^2}+\frac{2\dot{a}a^{(5)}}{35a^2}+\frac{a^{(6)}}{70a}\ , \tag{3.285}$$

where we have used the obvious notation $[a^{(n)} \equiv \frac{d^n}{dt^n}a]$. We have checked that (3.285) agrees with the third order DeWitt coefficient for fermions [34, 35],

$$\begin{aligned} -A_3(x) &= a_3(\xi = 1/4)\mathbb{I} - \Sigma^{[ab]}\Sigma^{[cd]}\left[\frac{R}{576}R_{ab\mu\nu}R_{cd}{}^{\mu\nu} + \frac{1}{720}R_{ab\mu\nu}{}^{;\mu}R_{cd\alpha}{}^{\nu;\alpha} + \frac{1}{120}R_{ab\mu\nu}R_{cd\alpha}{}^{\nu;\alpha\mu}\right. \\ &\left. + \frac{1}{180}R_{ab\mu\nu;\alpha}R_{cd}{}^{\mu\nu;\alpha} + \frac{1}{72}R^{\alpha\beta}R_{ab}{}^{\mu}{}_{\alpha}R_{cd\mu\beta} - \frac{1}{240}R^{\mu\nu\alpha\beta}R_{ab\mu\nu}R_{cd\alpha\beta}\right] \\ &+ \frac{1}{80}\Sigma^{[ab]}\Sigma^{[cd]}\Sigma^{[e,f]}R_{ab\mu\nu}R_{cd}{}^{\nu}{}_{\gamma}R_{ef}{}^{\gamma\mu} \, . \end{aligned} \tag{3.286}$$

We have also checked that this contribution is consistent with the purely sixth adiabatic order of the renormalized stress-energy tensor that has been reported in [134, 135].

Taking into account all this, it seems natural to argue that relations (3.283) and (3.284) are also valid for an arbitrary nth order, since both adiabatic and DeWitt-Schwinger methods provide a series expansion in which each contribution is univocally derived from some well-defined recursion relations using the first order terms as seeds for iteration. We have explicitly seen that the leading sixth order contributions agree, so it is very likely that higher order terms will agree as well. The calculation of the fourth and higher order DeWitt coefficients has been an elusive problem for a long time. The formal solution, given by a very involved recursion mechanism, was given in [34, 35]. To show the power of the adiabatic method for cosmological space-times, and also as an illustrative example, we have easily worked out the explicit form of the fourth DeWitt-Schwinger coefficient $a_4(x)$ using (3.376). It is given in Appendix I.

3.13 Extension of the equivalence to separate points

Finally, we would like to analyze the two-point functions, expanded up to a given adiabatic order, at separate points. The calculations are much more involved. We illustrate here explicitly the equivalence found at fourth adiabatic order for scalar fields. The adiabatic scheme provides the following result:

$$\begin{aligned} {}^{(4)}G_{Ad}((t,\vec{x}),(t,\vec{x}')) &= \frac{1}{2(2\pi)^3a^3}\int d^3\vec{k}\, e^{i\vec{k}\Delta\vec{x}}\left[\frac{1}{\omega} + \frac{(\frac{1}{6}-\xi)R}{2\omega^3} + \frac{m^2\dot{a}^2}{2a^2\omega^5} + \frac{m^2\ddot{a}}{4a\omega^5} - \frac{5m^4\dot{a}^2}{8a^2\omega^7} + (W^{-1})^{(4)}\right] \\ &= \frac{m}{4\pi^2 a|\Delta\vec{x}|}K_1(am|\Delta\vec{x}|) + \frac{(\frac{1}{6}-\xi)R}{8\pi^2}K_0(am|\Delta\vec{x}|) \\ &+ \frac{R}{288\pi^2}(ma|\Delta\vec{x}|)K_1(am|\Delta\vec{x}|) - \frac{H^2}{96\pi^2}(am|\Delta\vec{x}|)^2K_0(am|\Delta\vec{x}|) \\ &+ \frac{1}{2(2\pi)^3a^3}\int d^3\vec{k}\, e^{i\vec{k}\Delta\vec{x}}(W^{-1})^{(4)} \, , \end{aligned} \tag{3.287}$$

where

$$
\begin{aligned}
&\frac{1}{2(2\pi)^3a^3}\int d^3\vec{k}\, e^{i\vec{k}\Delta\vec{x}}(W^{-1})^{(4)} = \\
&\quad \frac{K_0(am|\Delta\vec{x}|)}{\pi^2}\left\{-\frac{7|\Delta\vec{x}|^4a^4H^4}{5760}-\frac{11m^2|\Delta\vec{x}|^4a^4H^2\ddot{a}}{5760a}-\frac{|\Delta\vec{x}|^2\xi aH^2\ddot{a}}{4a}+\frac{43|\Delta\vec{x}|^2a^2H^2\ddot{a}}{960a}\right.\\
&\qquad \left.+\frac{3|\Delta\vec{x}|^2\ddot{a}^2}{320a^2}-\frac{|\Delta\vec{x}|^2\xi\ddot{a}^2}{16a^2}+\frac{7|\Delta\vec{x}|^2}{960}H\frac{\dddot{a}}{a}-\frac{|\Delta\vec{x}|^2\xi H\dddot{a}}{16a}-\frac{|\Delta\vec{x}|^2a^2\ddddot{a}}{960a}\right\}\\
&\quad +\frac{K_1(am|\Delta\vec{x}|)}{\pi^2}\left\{H^4\left[\frac{|\Delta\vec{x}|a}{32m}-\frac{3\xi|\Delta\vec{x}|a}{8m}+\frac{9\xi^2|\Delta\vec{x}|a}{8m}-\frac{m|\Delta\vec{x}|^3a^3}{180}+\frac{m|\Delta\vec{x}|^3a^3\xi}{32}+\frac{m^3|\Delta\vec{x}|^5a^5}{4608}\right]\right.\\
&\qquad H^2\frac{\ddot{a}}{a}\left[\frac{m|\Delta\vec{x}|^3a^3}{2880}+\frac{m\xi|\Delta\vec{x}|^3a^3}{32}+\frac{29|\Delta\vec{x}|a}{240m}-\frac{17|\Delta\vec{x}|a\xi}{16m}+\frac{9|\Delta\vec{x}|a\xi^2}{4m}\right]\\
&\qquad \frac{\ddot{a}^2}{a^2}\left[\frac{3|\Delta\vec{x}|a}{160m}-\frac{-5\xi|\Delta\vec{x}|a}{16m}+\frac{9|\Delta\vec{x}|a\xi^2}{8m}+\frac{m|\Delta\vec{x}|^3a^3\xi}{640}\right]+\frac{\dddot{a}}{a}H\left[\frac{3|\Delta\vec{x}|a}{80m}+\frac{3|\Delta\vec{x}|a\xi}{16m}+\frac{m|\Delta\vec{x}|^3a^3}{480}\right]\\
&\qquad \left.+\frac{\ddddot{a}}{a}\left[-\frac{|\Delta\vec{x}|}{80m}+\frac{|\Delta\vec{x}|\xi}{16m}\right]\right\}.
\end{aligned}
$$

On the other hand, the DeWitt-Schwinger calculation provides (3.238). To compare it with the above result just expand it up to fourth adiabatic order (for the following identities we shall use the auxiliary parameter T to denote the number of time-derivatives that are present). Use

$$
g = 1+\frac{1}{3}R_{\alpha\beta}y^\alpha y^\beta+\frac{1}{6}R_{\alpha\beta;\gamma}y^\alpha y^\beta y^\gamma+\left[\frac{1}{18}R_{\alpha\beta}R_{\gamma\delta}-\frac{1}{90}R_{\lambda\alpha\beta}{}^{k}R^{\lambda}{}_{\gamma\delta k}+\frac{1}{20}R_{\alpha\beta;\gamma\delta}\right]y^\alpha y^\beta y^\gamma y^\delta+O(T^{-5})\,, \tag{3.291}
$$

and relations (3.242)-(3.244) including the fourth adiabatic contributions [51] (for simplicity we only show here, without loss of generality, the corresponding expressions for $\Delta t = 0$)

$$
y^0 = \frac{1}{2}Ha^2\Delta x^2+\frac{1}{144}a^4\Delta x^4HR+O(T^{-5})\,, \tag{3.292}
$$

$$
y^i = a\Delta x^i\left\{1+\frac{1}{6}a^2\Delta x^2H^2+\frac{1}{120}a^4\Delta x^4H^2\left[H^2+3\frac{\ddot{a}}{a}\right]\right\}+O(T^{-5})\,, \tag{3.293}
$$

$$
-2\sigma = \Delta x^2a^2+\frac{1}{12}a^4\Delta x^4H^2+\frac{1}{360}a^6\Delta x^6H^2\left[H^2+3\frac{\ddot{a}}{a}\right]+O(T^{-5})\,. \tag{3.294}
$$

Using all these auxiliary expressions we can prove that the adiabatic scheme generates the same two-point function as the DeWitt-Schwinger one up to fourth order in the derivatives of the metric,

$$
{}^{(4)}G_{DS}(x,x') = {}^{(4)}G_{Ad}(x,x')\,. \tag{3.295}
$$

We believe that one might extend this identity up to any order by induction. We note in passing that the result (3.295) implies the equivalence of the renormalized stress-energy tensor for $\xi \neq 1/6$ (see comments below (3.256)).

3.14 Summary and final comments

When a quantum field is coupled to a dynamical classical background, it gets excited, and undergoes a phase of particle creation. In this case new UV-divergent terms appear in the expectation values of its quadratic products, which must be appropriately removed to obtain a physical, finite quantity. In cosmological scenarios, adiabatic regularization provides an appropriate solution to this challenge: by means of an adiabatic expansion of the field modes in momenta, one can identify the covariant UV-divergent terms of the corresponding bilinear, and subtract them directly from the unrenormalized quantity. The background may be the expansion of the Universe itself, as in the case of inflation, or a classical homogeneous scalar field, as in the preheating phase. The adiabatic scheme can be applied in both situations, for both bosonic and fermionic species.

We began this chapter by presenting an extension of the adiabatic regularization method in order to find the renormalized spin-1/2 stress-energy tensor in a FLRW expanding universe. The main results are equations (3.72) and (3.73), which provide expressions that are simple and numerically easy to compute, once the quantum state is given. Then we illustrated this approach by briefly analyzing de Sitter space, with an assumed Bunch-Davies type vacuum state. The renormalized energy and pressure densities coincides with those predicted by symmetry arguments [126, 127]. We have also analyzed the renormalized stress-energy tensor in a purely radiation-dominated universe. In the latter case the early and late-time behavior of the renormalized stress-energy tensor can be worked out explicitly, irrespective of the specific form of the quantum state, and agree with those assumed by classical cosmology for radiation and cold matter, respectively.

After that, we developed the adiabatic regularization method for spin-1/2 fields in an expanding Universe when coupled to a classical background scalar field through a Yukawa interaction term. This is of particular physical interest in models of cosmological preheating. The results of this work are a natural generalization of the above, and broaden significantly the range of applicability of the adiabatic method. We have computed the adiabatic expansion of the spin-1/2 field modes up to 4th adiabatic order, and used it to obtain expressions for the renormalized expectation values of the stress-energy tensor $\langle T_{\mu\nu}\rangle_{ren}$ and the bilinear $\langle\bar{\psi}\psi\rangle_{ren}$. These quantities are fundamental ingredients in the study of the semiclassical equations of fermionic matter interacting with a background field, as they codify the backreaction effects from the created matter on the metric/background fields. Therefore, it is essential to develop an efficient renormalization scheme to correctly quantify the effects of this backreaction. All expressions obtained are generic, depending only on the background scalar field and scale factor time-dependent functions. This constitutes probably the major advantage of the adiabatic renormalization scheme. We leave the method prepared to perform numerical computations in future investigations for cosmological scenarios of interest.

We tested the overall theoretical construction by justifying the method in terms of renormalization of coupling constants, as well as by computing the conformal anomaly. Our calculation of the conformal anomaly with the Yukawa interaction has been proved to be fully consistent with the generic results obtained via the one-loop effective action. Therefore, by considering such a system, we have also improved our general understanding of quantum field theory in curved spacetimes.

A final goal of this chapter was to show the equivalence of the renormalized expectation values of the stress-energy tensor for spin-1/2 fields using both adiabatic and DeWitt-Schwinger methods. This is a very natural question since the adiabatic renormalization scheme for Dirac fields has been introduced very recently in the literature. The employed strategy to achieve our goal has led us to show the equivalence for scalar fields as well, in a simpler way to that used in [25, 47]. Moreover, we were naturally led to investigate the equivalence for the two-point function at coincidence for both DeWitt-Schwinger and adiabatic series expansion at any order. We have checked explicitly that the equality holds at sixth adiabatic order and we have argued that the equivalence must hold at an arbitrary order. This way, the adiabatic regularization method will offer a very efficient computational tool to evaluate the higher order DeWitt coefficients in FLRW space-times for both scalar and Dirac fields. This may be relevant to capture nonperturbative aspects of the effective action in cosmological space-times, as those found in [154–156, 158]. Finally, we would like to remark that these results suggest that the equality ${}^{(n)}G_{Ad}(x,x) = {}^{(n)}G_{DS}(x,x)$, $n = 0,2,4,6,...$ (and the analogue for Dirac fields) could even hold for separate points. This is actually supported by the fact that (3.255), (3.282), extended to separate points, coincide at least up to the fourth adiabatic order.

3.15 Appendices

A. Fermionic adiabatic expansion

We give here the third and fourth order contributions to $F(t)$ and $\omega(t)$ in (3.18) by solving the set of equations (3.19) order by order [we have $G^{(n)}(m) = F^{(n)}(-m)$].

$$\omega^{(3)} = 0\,, \tag{3.296}$$

$$F^{(3)} = i\left(\frac{65m^5\dot{a}^3}{64a^3\omega^8} - \frac{97m^3\dot{a}^3}{128a^3\omega^6} + \frac{m\dot{a}^3}{16a^3\omega^4} - \frac{19m^3\dot{a}\ddot{a}}{32a^2\omega^6} + \frac{m\dot{a}\ddot{a}}{4a^2\omega^4} + \frac{m\dddot{a}}{16a\omega^4}\right)\,, \tag{3.297}$$

and

$$\begin{aligned}\omega^{(4)} = & -\frac{1105m^8\dot{a}^4}{128a^4\omega^{11}} + \frac{337m^6\dot{a}^4}{32a^4\omega^9} - \frac{377m^4\dot{a}^4}{128a^4\omega^7} + \frac{3m^2\dot{a}^4}{32a^4\omega^5} + \frac{221m^6\dot{a}^2\ddot{a}}{32a^3\omega^9} - \frac{389m^4\dot{a}^2\ddot{a}}{64a^3\omega^7} + \frac{13m^2\dot{a}^2\ddot{a}}{16a^3\omega^5} - \frac{19m^4\ddot{a}^2}{32a^2\omega^7} \\ & + \frac{m^2\ddot{a}^2}{4a^2\omega^5} - \frac{7m^4\dot{a}\dddot{a}}{8a^2\omega^7} + \frac{15m^2\dot{a}\dddot{a}}{32a^2\omega^5} + \frac{m^2\ddddot{a}}{16a\omega^5}\,,\end{aligned} \tag{3.298}$$

$$\begin{aligned}F^{(4)} = & +\frac{2285m^8\dot{a}^4}{512a^4\omega^{12}} - \frac{565m^7\dot{a}^4}{128a^4\omega^{11}} - \frac{1263m^6\dot{a}^4}{256a^4\omega^{10}} + \frac{2611m^5\dot{a}^4}{512a^4\omega^9} + \frac{2371m^4\dot{a}^4}{2048a^4\omega^8} - \frac{333m^3\dot{a}^4}{256a^4\omega^7} - \frac{3m^2\dot{a}^4}{128a^4\omega^6} + \frac{m\dot{a}^4}{32a^4\omega^5} \\ & -\frac{457m^6\dot{a}^2\ddot{a}}{128a^3\omega^{10}} + \frac{113m^5\dot{a}^2\ddot{a}}{32a^3\omega^9} + \frac{725m^4\dot{a}^2\ddot{a}}{256a^3\omega^8} - \frac{749m^3\dot{a}^2\ddot{a}}{256a^3\omega^7} - \frac{19m^2\dot{a}^2\ddot{a}}{64a^3\omega^6} + \frac{11m\dot{a}^2\ddot{a}}{32a^3\omega^5} + \frac{41m^4\ddot{a}^2}{128a^2\omega^8} - \frac{5m^3\ddot{a}^2}{16a^2\omega^7} \\ & -\frac{17m^2\ddot{a}^2}{128a^2\omega^6} + \frac{m\ddot{a}^2}{8a^2\omega^5} + \frac{7m^4\dot{a}\dddot{a}}{16a^2\omega^8} - \frac{7m^3\dot{a}\dddot{a}}{16a^2\omega^7} - \frac{13m^2\dot{a}\dddot{a}}{64a^2\omega^6} + \frac{7m\dot{a}\dddot{a}}{32a^2\omega^5} - \frac{m^2\ddddot{a}}{32a\omega^6} + \frac{m\ddddot{a}}{32a\omega^5}\,.\end{aligned} \tag{3.299}$$

B. Useful formulas for a FLRW spacetime

In checking that the fourth order adiabatic subtraction terms (3.54) and (3.66) give the covariant result (3.69) we used the following results

$$R_{00} = 3\frac{\ddot{a}}{a}\,, \qquad R_{ij} = -a^2\left[2\frac{\dot{a}^2}{a^2} + \frac{\ddot{a}}{a}\right]\delta_{ij}\,, \tag{3.300}$$

$$R^2 = 36\left[\frac{\dot{a}^4}{a^4} + 2\frac{\dot{a}^2}{a^2}\frac{\ddot{a}}{a} + \frac{\ddot{a}^2}{a^2}\right]\,, \tag{3.301}$$

$$\Box R = 6\left[\frac{\ddot{a}^2}{a^2} + \frac{\ddddot{a}}{a} - 5\frac{\dot{a}^2}{a^2}\frac{\ddot{a}}{a} + 3\frac{\dot{a}}{a}\frac{\dddot{a}}{a}\right]\,, \tag{3.302}$$

$$R_{\mu\nu}R^{\mu\nu} = 12\left[\frac{\dot{a}^4}{a^4} + \frac{\ddot{a}^2}{a^2} + \frac{\dot{a}^2}{a^2}\frac{\ddot{a}}{a}\right]\,, \tag{3.303}$$

$$R_{;00} = 6\left[\frac{\ddddot{a}}{a} + \frac{\ddot{a}^2}{a^2} - 8\frac{\dot{a}^2}{a^2}\frac{\ddot{a}}{a} + 6\frac{\dot{a}^4}{a^4}\right]\,, \tag{3.304}$$

$$R_{;ij} = -6a^2\left[\frac{\dddot{a}}{a}\frac{\dot{a}}{a} + \frac{\ddot{a}}{a}\frac{\dot{a}^2}{a^2} - 2\frac{\dot{a}^4}{a^4}\right]\delta_{ij}\,. \tag{3.305}$$

C. Asymptotic analysis of the energy density and pressure for a free spin 1/2 field in a radiation dominated universe

In this Appendix we shall study the asymptotic properties of the stress tensor components in a radiation-dominated universe. The main results are collected and used in section 3.4, starting from formula (3.103). The large momentum behaviour of this tensor will give us a necessary and sufficient condition for its renormalizability, while the late/early-time behaviour will reproduce the classical results of physics of fluids for a matter/radiation-dominated universe.

Recall the general solution (3.102) for the modes in a radiation-dominated universe, and define for simplicity the following quantities,

$$
\begin{aligned}
g_k^I(t) &\equiv N\frac{W_{\kappa,\mu}(z)}{\sqrt{a(t)}}\,, && (3.306)\\
g_k^{II}(t) &\equiv N\frac{k}{2ma(t)^{3/2}}\left[W_{\kappa,\mu}(z)+\left(\kappa-\frac{3}{4}\right)W_{\kappa-1,\mu}(z)\right]. && (3.307)
\end{aligned}
$$

Expressions (3.42), (3.59), and (3.67), can be rewritten in terms of these independent solutions by doing a a Bogolubov-type rotation ($h_k^I \to E_k g_k^I + F_k g_k^{II*}$, $h_k^{II} \to E_k g_k^{II} - F_k g_k^{I*}$)

$$
\begin{aligned}
\rho_k &= \rho_k^D\left[|E|^2-|F|^2\right]+EF^*\rho_k^{ND}+E^*F\rho_k^{ND*}\,, && (3.308)\\
p_k &= p_k^D\left[|E|^2-|F|^2\right]+EF^*p_k^{ND}+E^*Fp_k^{ND*}\,, && (3.309)\\
\rho_k^D &= -3p_k^D-2m\left[|g_k^I|^2-|g_k^{II}|^2\right], && (3.310)
\end{aligned}
$$

where

$$
\begin{aligned}
\rho_k^D &= i\left[g_k^I\frac{\partial g_k^{I*}}{\partial t}+g_k^{II}\frac{\partial g_k^{II*}}{\partial t}-g_k^{I*}\frac{\partial g_k^I}{\partial t}-g_k^{II*}\frac{\partial g_k^{II}}{\partial t}\right], && (3.311)\\
\rho_k^{ND} &= -2i\left[g_k^I\frac{\partial g_k^{II}}{\partial t}-g_k^{II}\frac{\partial g_k^I}{\partial t}\right], && (3.312)\\
p_k^D &= -\frac{2k}{3a}\left[g_k^Ig_k^{II*}+g_k^{I*}g_k^{II}\right], && (3.313)\\
p_k^{ND} &= -\frac{2k}{3a}\left[(g^I)^2-(g^{II})^2\right]. && (3.314)
\end{aligned}
$$

The energy density and the pressure functions, expressed this way, show explicitly the dependence on the vacuum state.

Using the result (3.311), and derivative and functional properties of the Whittaker functions [8], one can find for the energy density,

$$\rho_k^D = -2m - \frac{4mx^2 e^{-\pi x^2}}{|z|^{3/2}} \times \left[|W_{\kappa,\frac{1}{4}}(z)|^2 - \left(\frac{1}{4}+x^4\right)|W_{\kappa-1,\frac{1}{4}}(z)|^2 \right] . \tag{3.315}$$

For large values of the momenta, $k^2/a^2 >> m^2$, the Whittaker function can be very well approximated by [130]

$$W_{\kappa,1/4}(z) = \frac{\sqrt{\pi}z^{1/4}}{\Gamma\left(\frac{3}{4}-\kappa\right)}\left\{\cos(2\sqrt{\kappa z})\left[1-\frac{M_2(z)}{\kappa}+\frac{M_4(z)}{\kappa^2}-\left(\frac{M_1(z)}{\kappa}-\frac{M_3(z)}{\kappa^2}\right)\frac{\Gamma(3/4-\kappa)}{\Gamma(1/4-\kappa)}\right]\right. \\ \left. - \frac{\sin(2\sqrt{\kappa z})}{\sqrt{\kappa}}\left[M_1(z)-\frac{M_3(z)}{\kappa}+\frac{M_5(z)}{\kappa^2}+\left(1-\frac{M_2(z)}{\kappa}+\frac{M_4(z)}{\kappa^2}\right)\frac{\Gamma(3/4-\kappa)}{\Gamma(1/4-\kappa)}\right]\right\}+O(|\kappa|^{-3}) , \tag{3.316}$$

where $M_n(z)$ are a set of polynomials that satisfy

$$M_1(z) = -\frac{z^{3/2}}{12} , \tag{3.317}$$

$$M_2(z) = -\frac{z}{16}\left(1-\frac{z^2}{18}\right) , \tag{3.318}$$

$$M_3(z) = -\frac{z^{1/2}}{32}+\frac{z^{5/2}}{120}-\frac{z^{9/2}}{10368} , \tag{3.319}$$

$$M_4(z) = -\frac{1}{128}+\frac{19z^2}{1536}-\frac{11z^4}{23040}+\frac{z^6}{497664} , \tag{3.320}$$

$$M_5(z) = \frac{z^{3/2}(2721600-291924z^2+3528z^4-7z^6)}{209018880} . \tag{3.321}$$

After a long calculation one can find that

$$|W_{\kappa,\frac{1}{4}}(z)|^2 - \left(\frac{1}{4}+x^4\right)|W_{\kappa-1,\frac{1}{4}}(z)|^2 = \left(\frac{z}{x}-\frac{z^{3/2}}{2x^2}+\frac{z^2}{8x^3}-\frac{z(1+z^2)}{128x^5}+O(x^{-7})\right)e^{\pi x^2} , \tag{3.322}$$

so that, taking the change $x=\sqrt{\frac{t}{2m}(w^2-m^2)}$, one gets

$$\rho_k^D = \left[-2w+\frac{m^2}{16t^2w^3}+O(w^{-5})\right] . \tag{3.323}$$

From this expression it is easy to see that we recover those terms of zeroth and second adiabatic order found in Eqs. (3.45)-(3.50) for a radiation dominated universe. These contributions give the divergences of the stress-energy tensor. Additionally, the no-diagonal terms are shown to be

$$\rho_k^{ND} = O(\omega^{-1}) . \tag{3.324}$$

On the other hand, taking (3.313) (or more easily (3.310)) we can find for the pressure

$$\begin{aligned} p_k^D &= -\frac{4m}{3} - \frac{4mx^2 e^{-\pi x^2}}{3|z|^{3/2}} \\ &\times \left[\left(1-\frac{z}{x^2}\right)|W_{\kappa,\frac{1}{4}}(z)|^2 - \left(\frac{1}{4}+x^4\right)|W_{\kappa-1,\frac{1}{4}}(z)|^2\right] \\ &= -\left[\frac{2}{3}\omega - \frac{2m^2}{3w} - \frac{m^2}{48t^2\omega^3} + O(\omega^{-7})\right] , \end{aligned} \tag{3.325}$$

which also agrees with the divergences found in (3.63)-(3.64). Additionally,

$$p_k^{ND} = O(\omega^{-1}) . \tag{3.326}$$

The choice of the parameters E_k and F_k (the choice of the vacuum state) is determined by imposing some initial condition at a given instant of time, t_0. This choice must be in such a way that leave the stress energy tensor without divergences. According to (3.323), (3.325) and (3.53), (3.61) respectively, the stress energy tensor renormalizability imposes a natural constraint on the vacuum state (recall (3.308) and (3.309)),

$$|E_k|^2 - |F_k|^2 = 1 + O(w^{-5}) . \tag{3.327}$$

This means that $E_k = 1 + O(w^{-5})$ and $F_k = O(w^{-5/2})$, which makes $E_k F_k^* = O(w^{-5/2})$, and it is enough for the no-diagonal terms, (3.324) and (3.326), to not to give new divergences.

Let's focus now on the stress-energy tensor for late times in the expansion of the universe. Taking $t >> m^{-1}$, equation (3.315) behaves as

$$\rho_k^D = -2m - \frac{4mx^2}{z} + \frac{4mx^4}{z^2} + \dots , \tag{3.328}$$

while

$$\rho_k^{(0)} + \rho_k^{(2)} + \rho_k^{(4)} = \rho_k^D + O(z^{-7}) , \tag{3.329}$$

so we may state, recalling (3.53),

$$\langle T_{00}\rangle_{ren}(t >> m^{-1}) = \frac{1}{2\pi^2 a^3}\left[\int_0^\infty dk k^2 2m\left[1-(|E_k|^2-|F_k|^2)\right] + O(z^{-1})\right] . \tag{3.330}$$

Similarly, one can study the late times behaviour of the pressure (3.325), and find

$$p_k^D = -\frac{8mx^2}{3z} + \frac{16mx^4}{3z^2} - \frac{2mx^2(-1+8x^4)}{z^3} + \dots , \tag{3.331}$$

while the corresponding adiabatic subtractions are

$$p_k^{(0)} + p_k^{(2)} + p_k^{(4)} = p_k^D + O(z^{-7}) . \tag{3.332}$$

Again, following (3.61) we find

$$\langle T_{ii}\rangle_{ren}(t>>m^{-1}) = \frac{1}{2\pi^2 a}\left[\int_0^\infty dk k^2 \frac{8mx^2}{3z}\left[1-(|E_k|^2-|F_k|^2)\right]+O(z^{-2})\right] . \quad (3.333)$$

This time, the dominant contribution to the total pressure, $p \equiv \langle T_{ii}\rangle_{ren}/a^2$, decays with time. Basically, equations (3.330) and (3.333) tell us that in a radiation dominated expansion of the universe, a spin 1/2 field tend to behave as a source of cold matter in cosmology. This may be useful to analyze in detail the phase transition from radiation to matter dominated universes, in the standard cosmology.

On the other hand, at early times $t << m^{-1}$, (3.315) reads [we analyze only the large momentum behaviour since it is in this case where any problem with divergences might arise]

$$\rho_k^D = i\frac{4m\pi e^{-\pi x^2}}{\sqrt{z}}\left[\frac{(-1)^{1/4}}{\Gamma(ix^2)\Gamma(1/2-ix^2)}+\frac{(-1)^{3/4}}{\Gamma(-ix^2)\Gamma(1/2+ix^2)}\right]+O(z^0) \quad (3.334)$$

$$= \frac{1}{\sqrt{z}}\left[-4mx+\frac{m}{32x^3}+\frac{21m}{8192x^7}+O(x^{-11})\right]+O(z^0) , \quad (3.335)$$

and just as in the late-time case we obtain

$$\rho_k^{(0)}+\rho_k^{(2)}+\rho_k^{(4)} = \frac{1}{\sqrt{z}}\left[-4mx+\frac{m}{32x^3}+\frac{21m}{8192x^7}\right]+O(z^{1/2}) , \quad (3.336)$$

so at early times,

$$\langle T_{00}\rangle_{ren}(t<<m^{-1}) \approx \frac{1}{2\pi^2 a^3}\left[\int_0^\infty dk k^2 \frac{4mx}{\sqrt{z}}\left[1-(|E_k|^2-|F_k|^2)\right]+O(z^0)\right] . \quad (3.337)$$

Finally, if one tries to do the same calculation to the pressure (3.325), one finds just the same results for (3.335) and (3.336) but with a factor 1/3, so recovering this way the equation of state for classical radiation in cosmology.

D. Adiabatic expansion with Yukawa coupling. Explicit expressions

In this appendix, we provide the terms of the adiabatic expansion of the spin-1/2 field modes up to fourth order, discussed in section 3.6. Although the first and second order terms have already been written in there, we copy them here for convenience. As introduced in equations (3.131) and (3.132), the adiabatic expansion takes the form

$$h_k^I(t) = \sqrt{\frac{\omega+m}{2\omega}}e^{-i\int^t(\omega+\omega^{(1)}+\omega^{(2)}+\omega^{(3)}+\omega^{(4)}+\dots)dt'}(1+F^{(1)}+F^{(2)}+F^{(3)}+F^{(4)}+\dots) ,$$

$$h_k^{II}(t) = \sqrt{\frac{\omega-m}{2\omega}}e^{-i\int^t(\omega+\omega^{(1)}+\omega^{(2)}+\omega^{(3)}+\omega^{(4)}+\dots)dt'}(1+G^{(1)}+G^{(2)}+G^{(3)}+G^{(4)}+\dots) \quad (3.338)$$

The terms $G^{(n)}$ can be obtained from $F^{(n)}$ with the relation $G^{(n)}(m,s) = F^{(n)}(-m,-s)$, so we do not explicitly write them here. We denote by $f_x^{(n)}$ and $f_y^{(n)}$ to the real and imaginary parts of $F^{(n)}$ respectively, so that $F^{(n)} = f_x^{(n)} + i f_y^{(n)}$.

The first order terms are:

$$f_x^{(1)} = \frac{s}{2\omega} - \frac{ms}{2\omega^2}\,, \tag{3.339}$$
$$f_y^{(1)} = -\frac{m\dot{a}}{4\omega^2 a}\,, \tag{3.340}$$
$$\omega^{(1)} = \frac{ms}{\omega}\,. \tag{3.341}$$

The second order terms are:

$$f_x^{(2)} = \frac{m^2\ddot{a}}{8a\omega^4} - \frac{m\ddot{a}}{8a\omega^3} - \frac{5m^4\dot{a}^2}{16a^2\omega^6} + \frac{5m^3\dot{a}^2}{16a^2\omega^5} + \frac{3m^2\dot{a}^2}{32a^2\omega^4} - \frac{m\dot{a}^2}{8a^2\omega^3} + \frac{5m^2s^2}{8\omega^4} - \frac{ms^2}{2\omega^3} - \frac{s^2}{8\omega^2}\,, \tag{3.342}$$
$$f_y^{(2)} = \frac{5m^2s\dot{a}}{8a\omega^4} - \frac{s\dot{a}}{4a\omega^2} - \frac{\dot{s}}{4\omega^2}\,, \tag{3.343}$$
$$\omega^{(2)} = -\frac{m^2s^2}{2\omega^3} + \frac{s^2}{2\omega} + \frac{5m^4\dot{a}^2}{8a^2\omega^5} - \frac{3m^2\dot{a}^2}{8a^2\omega^3} - \frac{m^2\ddot{a}}{4a\omega^3}\,. \tag{3.344}$$

The third order terms are:

$$\begin{aligned} f_x^{(3)} = & -\frac{15m^3s^3}{16\omega^6} + \frac{11m^2s^3}{16\omega^5} + \frac{7ms^3}{16\omega^4} - \frac{3s^3}{16\omega^3} + \frac{65m^5s\dot{a}^2}{32a^2\omega^8} - \frac{15m^4s\dot{a}^2}{8a^2\omega^7} - \frac{97m^3s\dot{a}^2}{64a^2\omega^6} + \frac{93m^2s\dot{a}^2}{64a^2\omega^5} \\ & + \frac{ms\dot{a}^2}{8a^2\omega^4} - \frac{s\dot{a}^2}{8a^2\omega^3} - \frac{5m^3\dot{a}\dot{s}}{8a\omega^6} + \frac{5m^2\dot{a}\dot{s}}{8a\omega^5} + \frac{5m\dot{a}\dot{s}}{16a\omega^4} - \frac{3\dot{a}\dot{s}}{8a\omega^3} - \frac{9m^3s\ddot{a}}{16a\omega^6} + \frac{m^2s\ddot{a}}{2a\omega^5} + \frac{3ms\ddot{a}}{16a\omega^4} \\ & - \frac{s\ddot{a}}{8a\omega^3} + \frac{m\ddot{s}}{8\omega^4} - \frac{\ddot{s}}{8\omega^3}\,, \end{aligned} \tag{3.345}$$
$$\begin{aligned} f_y^{(3)} = & -\frac{45m^3s^2\dot{a}}{32a\omega^6} + \frac{31ms^2\dot{a}}{32a\omega^4} + \frac{65m^5\dot{a}^3}{64a^3\omega^8} - \frac{97m^3\dot{a}^3}{128a^3\omega^6} + \frac{m\dot{a}^3}{16a^3\omega^4} + \frac{5ms\dot{s}}{8\omega^4} - \frac{19m^3\dot{a}\ddot{a}}{32a^2\omega^6} + \frac{m\dot{a}\ddot{a}}{4a^2\omega^4} \\ & + \frac{ma^{(3)}}{16a\omega^4}\,, \end{aligned} \tag{3.346}$$
$$\omega^{(3)} = \frac{m^3s^3}{2\omega^5} - \frac{ms^3}{2\omega^3} - \frac{25m^5s\dot{a}^2}{8a^2\omega^7} + \frac{13m^3s\dot{a}^2}{4a^2\omega^5} - \frac{ms\dot{a}^2}{2a^2\omega^3} + \frac{5m^3\dot{a}\dot{s}}{4a\omega^5} - \frac{7m\dot{a}\dot{s}}{8a\omega^3} + \frac{3m^3s\ddot{a}}{4a\omega^5} - \frac{3ms\ddot{a}}{8a\omega^3} - \frac{m\ddot{s}}{4\omega^3}\,. \tag{3.347}$$

Finally, the fourth-order terms are:

$$\begin{aligned}
f_x^{(4)} = {} & \frac{2285\dot{a}^4m^8}{512a^4\omega^{12}} - \frac{565\dot{a}^4m^7}{128a^4\omega^{11}} - \frac{1263\dot{a}^4m^6}{256a^4\omega^{10}} - \frac{1105s^2\dot{a}^2m^6}{128a^2\omega^{10}} - \frac{457\dot{a}^2\ddot{a}m^6}{128a^3\omega^{10}} + \frac{2611\dot{a}^4m^5}{512a^4\omega^9} + \frac{965s^2\dot{a}^2m}{128a^2\omega^9} \\
& + \frac{113\dot{a}^2\ddot{a}m^5}{32a^3\omega^9} + \frac{2371\dot{a}^4m^4}{2048a^4\omega^8} + \frac{2441s^2\dot{a}^2m^4}{256a^2\omega^8} + \frac{41\ddot{a}^2m^4}{128a^2\omega^8} + \frac{65s\dot{a}\dot{s}m^4}{16a\omega^8} + \frac{725\dot{a}^2\ddot{a}m^4}{256a^3\omega^8} + \frac{117s^2\ddot{a}m^4}{64a\omega^8} \\
& + \frac{7\dot{a}a^{(3)}m^4}{16a^2\omega^8} + \frac{195s^4m^4}{128\omega^8} - \frac{333\dot{a}^4m^3}{256a^4\omega^7} - \frac{1049s^2\dot{a}^2m^3}{128a^2\omega^7} - \frac{5\ddot{a}^2m^3}{16a^2\omega^7} - \frac{15s\dot{a}\dot{s}m^3}{4a\omega^7} - \frac{749\dot{a}^2\ddot{a}m^3}{256a^3\omega^7} \\
& - \frac{97s^2\ddot{a}m^3}{64a\omega^7} - \frac{7\dot{a}a^{(3)}m^3}{16a^2\omega^7} - \frac{17s^4m^3}{16\omega^7} - \frac{3\dot{a}^4m^2}{128a^4\omega^6} - \frac{561s^2\dot{a}^2m^2}{256a^2\omega^6} - \frac{5\dot{s}^2m^2}{16\omega^6} - \frac{17\ddot{a}^2m^2}{128a^2\omega^6} \\
& - \frac{95s\dot{a}\dot{s}m^2}{32a\omega^6} - \frac{19\dot{a}^2\ddot{a}m^2}{64a^3\omega^6} - \frac{73s^2\ddot{a}m^2}{64a\omega^6} - \frac{9s\ddot{s}m^2}{16\omega^6} - \frac{13\dot{a}a^{(3)}m^2}{64a^2\omega^6} - \frac{a^{(4)}m^2}{32a\omega^6} - \frac{71s^4m^2}{64\omega^6} \\
& + \frac{\dot{a}^4m}{32a^4\omega^5} + \frac{111s^2\dot{a}^2m}{64a^2\omega^5} + \frac{5\dot{s}^2m}{16\omega^5} + \frac{\ddot{a}^2m}{8a^2\omega^5} + \frac{89s\dot{a}\dot{s}m}{32a\omega^5} + \frac{11\dot{a}^2\ddot{a}m}{32a^3\omega^5} + \frac{49s^2\ddot{a}m}{64a\omega^5} + \frac{s\ddot{s}m}{2\omega^5} \\
& + \frac{7\dot{a}a^{(3)}m}{32a^2\omega^5} + \frac{a^{(4)}m}{32a\omega^5} + \frac{9s^4m}{16\omega^5} + \frac{s^2\dot{a}^2}{32a^2\omega^4} - \frac{\dot{s}^2}{32\omega^4} + \frac{s\dot{a}\dot{s}}{8a\omega^4} + \frac{s^2\ddot{a}}{16a\omega^4} + \frac{s\ddot{s}}{16\omega^4} + \frac{11s^4}{128\omega^4}\,, \qquad (3.348)
\end{aligned}$$

$$\begin{aligned}
f_y^{(4)} = {} & \frac{195m^4s^3\dot{a}}{64a\omega^8} - \frac{187m^2s^3\dot{a}}{64a\omega^6} + \frac{11s^3\dot{a}}{32a\omega^4} - \frac{1105m^6s\dot{a}^3}{128a^3\omega^{10}} + \frac{2571m^4s\dot{a}^3}{256a^3\omega^8} - \frac{329m^2s\dot{a}^3}{128a^3\omega^6} + \frac{s\dot{a}^3}{16a^3\omega^4} \\
& - \frac{45m^2s^2\dot{s}}{32\omega^6} + \frac{11s^2\dot{s}}{32\omega^4} + \frac{195m^4\dot{a}^2\dot{s}}{64a^2\omega^8} - \frac{367m^2\dot{a}^2\dot{s}}{128a^2\omega^6} + \frac{7\dot{a}^2\dot{s}}{16a^2\omega^4} + \frac{247m^4s\dot{a}\ddot{a}}{64a^2\omega^8} - \frac{187m^2s\dot{a}\ddot{a}}{64a^2\omega^6} + \frac{s\dot{a}\ddot{a}}{4a^2\omega^4} \\
& - \frac{19m^2\dot{s}\ddot{a}}{32a\omega^6} + \frac{\dot{s}\ddot{a}}{4a\omega^4} - \frac{19m^2\dot{a}\ddot{s}}{32a\omega^6} + \frac{3\dot{a}\ddot{s}}{8a\omega^4} - \frac{9m^2sa^{(3)}}{32a\omega^6} + \frac{sa^{(3)}}{16a\omega^4} + \frac{s^{(3)}}{16\omega^4}\,, \qquad (3.349)
\end{aligned}$$

$$\begin{aligned}
\omega^{(4)} = {} & -\frac{5m^4s^4}{8\omega^7} + \frac{3m^2s^4}{4\omega^5} - \frac{s^4}{8\omega^3} + \frac{175m^6s^2\dot{a}^2}{16a^2\omega^9} - \frac{245m^4s^2\dot{a}^2}{16a^2\omega^7} + \frac{79m^2s^2\dot{a}^2}{16a^2\omega^5} - \frac{s^2\dot{a}^2}{8a^2\omega^3} - \frac{1105m^8\dot{a}^4}{128a^4\omega^{11}} \\
& + \frac{337m^6\dot{a}^4}{32a^4\omega^9} - \frac{377m^4\dot{a}^4}{128a^4\omega^7} + \frac{3m^2\dot{a}^4}{32a^4\omega^5} - \frac{25m^4s\dot{a}\dot{s}}{4a\omega^7} + \frac{23m^2s\dot{a}\dot{s}}{4a\omega^5} - \frac{3s\dot{a}\dot{s}}{8a\omega^3} + \frac{5m^2\dot{s}^2}{8\omega^5} - \frac{15m^4s^2\ddot{a}}{8a\omega^7} \\
& + \frac{25m^2s^2\ddot{a}}{16a\omega^5} - \frac{s^2\ddot{a}}{8a\omega^3} + \frac{221m^6\dot{a}^2\ddot{a}}{32a^3\omega^9} - \frac{389m^4\dot{a}^2\ddot{a}}{64a^3\omega^7} + \frac{13m^2\dot{a}^2\ddot{a}}{16a^3\omega^5} - \frac{19m^4\ddot{a}^2}{32a^2\omega^7} + \frac{m^2\ddot{a}^2}{4a^2\omega^5} + \frac{3m^2s\ddot{s}}{4\omega^5} \\
& - \frac{s\ddot{s}}{8\omega^3} - \frac{7m^4\dot{a}a^{(3)}}{8a^2\omega^7} + \frac{15m^2\dot{a}a^{(3)}}{32a^2\omega^5} + \frac{m^2a^{(4)}}{16a\omega^5}\,. \qquad (3.350)
\end{aligned}$$

E. A simple realization of the Yukawa coupling

In this Appendix we consider a simple mathematical example to illustrate how the adiabatic method works. This is complementary to the results presented in 3.7. We compute the bilinear $\langle\bar{\psi}\psi\rangle_{ren}$ of a Dirac field, coupled to a background scalar field evolving in Minkowski spacetime ($a(t) = 1$) as

$$s(t) = g_Y\Phi(t) = \mu/t\ . \qquad (3.351)$$

For convenience, we have absorbed the Yukawa coupling g_Y in the dimensionless constant μ. To avoid the mathematical instability at $t \to 0$, we will only consider times in the range $-\infty < t < 0$. This model has three convenient aspects which simplify significantly the analysis. First of all, the field mode equations (3.127) have an analytical solution in terms of the well-known Whitakker functions, so we do not have to solve the equation numerically. Second, at time $t \to -\infty$ we have $s, \dot{s} \cdots \to 0$, so that the system is adiabatic initially, and there is no ambiguity when imposing initial conditions to the field modes. And third, as we shall see, the system behaves in such a way that, as long as we are well before the instability, $\langle \bar{\psi}\psi \rangle_{ren}$ can be approximated by the fourth order in its adiabatic expansion, giving a final renormalized bilinear that can be easily integrated.

It is useful to define a new dimensionless time $z \equiv mt$ and momenta $\kappa \equiv k/m$. The field equations (3.127) for h_k^I and h_k^{II} in terms of these variables become

$$h_k^{II} = \frac{i}{\kappa}\left[\frac{\partial h_k^I}{\partial z} + i\left(1+\frac{\mu}{z}\right)h_k^I\right] , \quad h_k^I = \frac{i}{\kappa}\left[\frac{\partial h_k^{II}}{\partial z} - i\left(1+\frac{\mu}{z}\right)h_k^{II}\right] , \tag{3.352}$$

and from these, we obtain the second-order uncoupled equations

$$\frac{d^2 h_k^I}{dz^2} + \left(1+\kappa^2+\frac{2\mu}{z}+\frac{\mu(\mu-i)}{z^2}\right)h_k^I = 0 , \quad \frac{d^2 h_k^{II}}{dz^2} + \left(1+\kappa^2+\frac{2\mu}{z}+\frac{\mu(\mu+i)}{z^2}\right)h_k^{II} = 0 . \tag{3.353}$$

Let us also define a dimensionless frequency $\omega_\kappa \equiv \sqrt{\kappa^2+1}$, so that $\omega = \sqrt{k^2+m^2} = m\omega_\kappa$. The general solution for $h_k^I(t)$ is a linear combination of the first and second kind Whittaker functions $M_{\alpha,\lambda_1}(2i\omega_\kappa t)$ and $W_{\alpha,\lambda_1}(2i\omega_\kappa t)$, where $\alpha \equiv \frac{-i\mu}{\sqrt{\kappa^2+1}}$ and $\lambda_1 \equiv -\frac{1}{2}i(2\mu - i)$. The solution for $h_k^{II}(t)$ is similar, with the change $\lambda_1 \to \lambda_2 \equiv -\frac{1}{2}i(2\mu+i)$, so we have

$$\begin{aligned} h_k^I &= A_k^I M_{\alpha,\lambda_1}(2i\omega_\kappa z) + B_k^I W_{\alpha,\lambda_1}(2i\omega_\kappa z) , \\ h_k^{II} &= A_k^{II} M_{\alpha,\lambda_2}(2i\omega_\kappa z) + B_k^{II} W_{\alpha,\lambda_2}(2i\omega_\kappa z) . \end{aligned} \tag{3.354}$$

Note that h_k^I and h_k^{II} must obey the constraint (3.128), so there is only one degree of freedom in the fermion solution, which is determined when imposing the initial conditions. To fix the constants in the linear combinations, we impose the adiabatic behaviour (3.130) at $z \to -\infty$, getting

$$A_k^I = A_k^{II} = 0 , \quad B_k^I = \sqrt{\frac{\omega_\kappa+1}{2\omega_\kappa}}\, e^{\frac{\mu\pi}{2\omega_\kappa}} , \quad B_k^{II} = \sqrt{\frac{\omega_\kappa-1}{2\omega_\kappa}}\, e^{\frac{\mu\pi}{2\omega_\kappa}} . \tag{3.355}$$

The final solution is then

$$h_k^I = \sqrt{\frac{\omega_\kappa+1}{2\omega_\kappa}}\, e^{\frac{\mu\pi}{2\omega_\kappa}} W_{\alpha,\lambda_1}(2i\omega_\kappa z) , \quad h_k^{II} = \sqrt{\frac{\omega_\kappa-1}{2\omega_\kappa}}\, e^{\frac{\mu\pi}{2\omega_\kappa}} W_{\alpha,\lambda_2}(2i\omega_\kappa z) . \tag{3.356}$$

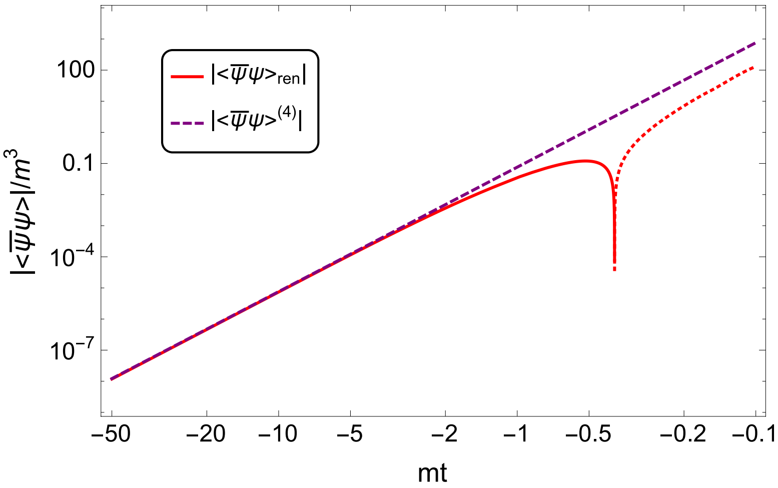

Figure 3.1: The red line shows $\frac{1}{m^3}|\langle\bar{\psi}\psi\rangle_{ren}|$ as a function of time for $\mu = 1$, given by equation (3.357). For $mt \lesssim -0.4$ we have $\langle\bar{\psi}\psi\rangle_{ren} < 0$ (red continuous line), while for $mt \gtrsim -0.4$ we have $\langle\bar{\psi}\psi\rangle_{ren} > 0$ (red dotted line). The purple dashed line shows the corresponding approximation at 4th adiabatic order, given in equations (3.360).

The renormalized expectation value $\langle\bar{\psi}\psi\rangle_{ren}$ is given, from (3.171), by

$$\langle\bar{\psi}\psi\rangle_{ren} = \frac{-m^3}{\pi^2}\int_0^\infty d\kappa\kappa^2 \left(|h_k^I|^2 - |h_k^{II}|^2 - \frac{1}{\omega_\kappa} - \frac{\mu}{\omega_\kappa z} + \frac{\mu}{\omega_\kappa^3 z} + \frac{3\mu^2}{2\omega_\kappa^3 z^2} - \frac{3\mu^2}{2\omega_\kappa^5 z^2} + \frac{\mu+\mu^3}{2\omega_\kappa^3 z^3} - \frac{\mu+6\mu^3}{2\omega_\kappa^5 z^3} + \frac{5\mu^3}{2\omega_\kappa^7 z^3}\right), \tag{3.357}$$

where we have that the adiabatic contributions of order n go as $\langle\bar{\psi}\psi\rangle^{(n)} \propto c_n(\mu,\kappa)z^{-n}$, with c_n time-independent functions of μ and κ. The above integral is finite, as one can easily check from the asymptotic expansion of the Whittaker function $W_{\alpha,\lambda}(x)$.

We can compute analytically the leading term at $z \to -\infty$ by performing the adiabatic expansion of $\langle\bar{\psi}\psi\rangle$ up to fourth order, and subtracting from it the 0th, 1st, 2nd, and 3rd orders. Therefore, the leading behaviour at very early times is

$$\begin{aligned}\langle\bar{\psi}\psi\rangle_{ren} \sim \langle\bar{\psi}\psi\rangle^{(4)} &\equiv -\frac{1}{\pi^2 a^3}\int_0^\infty dk k^2\left(\left(|h_k^I|^2\right)^{(4)} - \left(|h_k^{II}|^2\right)^{(4)}\right) \\ &= -\frac{1}{\pi^2 a^3}\int_0^\infty dk k^2\left(\frac{(\omega-m)}{2\omega}[G^{(4)}+G^{(4)*}+G^{(1)}G^{(3)*}+G^{(1)*}G^{(3)}+|G^{(2)}|^2]\right. \\ &\left. -\frac{(\omega+m)}{2\omega}[F^{(4)}+F^{(4)*}+F^{(1)}F^{(3)*}+F^{(1)*}F^{(3)}+|F^{(2)}|^2]\right).\end{aligned} \tag{3.358}$$

Computing the integral, we finally get

$$\begin{aligned}\langle\bar{\psi}\psi\rangle^{(4)} &= -\frac{a^{(4)}}{80\pi^2 am}+\frac{\dot{a}^2\ddot{a}}{60\pi^2a^3m}-\frac{\dot{a}^2s^2}{8\pi^2a^2m}-\frac{\ddot{a}^2}{80\pi^2a^2m}-\frac{3\dot{a}a^{(3)}}{80\pi^2a^2m}-\frac{s^2\ddot{a}}{8\pi^2am}\\ &\quad -\frac{3\dot{a}\dot{s}s}{4\pi^2am}-\frac{s^4}{8\pi^2m}-\frac{s\ddot{s}}{4\pi^2m}-\frac{\dot{s}^2}{8\pi^2m}\ .\end{aligned} \tag{3.359}$$

Substituting (3.351) in this expression, and setting $a=1$, we finally obtain

$$\langle\bar{\psi}\psi\rangle^{(4)} = -\frac{m^3\mu^2(\mu^2+5)}{8\pi^2z^4}\ , \tag{3.360}$$

where we have written the solution in terms of z. In Figure 3.1 we show $\frac{1}{m^3}|\langle\bar{\psi}\psi\rangle|_{ren}$ as a function of time, comparing the exact result (3.357) with the approximation (3.360). At very early times $z\to-\infty$ we have, as expected, $\langle\bar{\psi}\psi\rangle_{ren}\sim 0$. We observe that the approximation holds quite well, except when the instability is approached.

F. Scalar field with a Yukawa-type coupling

In this appendix we compute the conformal anomaly of a quantized real scalar field ϕ, coupled to another background scalar Φ with a Yukawa-type interaction. This result is used in Section 3.9. The interaction term can be chosen of the form $g\Phi\phi^2$ or $h^2\Phi^2\phi^2$. Although the adiabatic regularization can be equally applied in both cases, we will focus on the latter case, since the coupling constant h^2 is dimensionless and the classical theory inherits the conformal invariance. Therefore, the action functional of the scalar matter field is given by

$$S_m = \int d^4x\sqrt{-g}\frac{1}{2}(g^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi - m^2\phi^2 - \xi R\phi^2 - h^2\Phi^2\phi^2)\ . \tag{3.361}$$

As before, the scalar field lives in a spatially flat FLRW metric $ds^2 = dt^2 - a^2(t)d\vec{x}^2$, and we assume that the external field is homogeneous $\Phi=\Phi(t)$. In this case, the equation of motion is

$$(\Box + m^2 + s^2(t) + \xi R)\phi = 0\ , \tag{3.362}$$

where we have introduced the notation $s(t)\equiv h\Phi(t)$, similar to the one used for the spin-1/2 field in the main text. The quantized field is expanded in Fourier modes as

$$\phi(x) = \frac{1}{\sqrt{2(2\pi a^3)}}\int d^3\vec{k}[A_{\vec{k}}f_{\vec{k}}(x) + A^\dagger_{\vec{k}}f^*_{\vec{k}}(x)]\ , \tag{3.363}$$

where $f_{\vec{k}}(x) = e^{i\vec{k}\vec{x}}h_k(t)$, and $A^\dagger_{\vec{k}}$ and $A_{\vec{k}}$ are the usual creation and annihilation operators. Substituting (3.363) into (3.362) we find

$$\frac{d^2}{dt^2}h_k(t) + \left[\omega_k^2(t) + s^2(t) + \sigma(t)\right]h_k(t) = 0\ , \tag{3.364}$$

where $\sigma(t) = (6\xi - \frac{3}{4})(\frac{\dot{a}^2}{a^2}) + (6\xi - \frac{3}{2})(\frac{\ddot{a}}{a})$, and $\omega_k(t) = \sqrt{\frac{k^2}{a(t)^2} + m^2}$. The adiabatic expansion for the scalar field modes is based on the usual WKB ansatz

$$h_k(t) = \frac{1}{\sqrt{W_k}} e^{-i\int^t W_k(t')dt'} \ , \quad W_k(t) = \omega_k + \omega^{(1)} + \omega^{(2)} + \cdots \ , \tag{3.365}$$

which solves the Wronskian condition $h_k \dot{h}_k^* - h_k^* \dot{h}_k = 2i$. One can substitute the ansatz into equation (3.364), and solve order by order to obtain the different terms of the expansion. The function $W_k(t)$ obeys the differential equation

$$W_k^4 = (\omega^2 + s^2 + \sigma) W_k^2 + \frac{3}{4}\dot{W}_k^2 - \frac{1}{2}\ddot{W}_k^2 W_k \ . \tag{3.366}$$

Note that here, $s(t) \equiv h\Phi$ is assumed of adiabatic order 1 as in the fermionic case. One obtains systematically $\omega^{(odd)} = 0$ for all terms of odd order in the expansion. At second adiabatic order one gets

$$\begin{aligned} \omega^{(2)} &= \frac{1}{2\omega}(s^2 + \sigma) + \frac{3\dot{\omega}^2}{8\omega^3} - \frac{\ddot{\omega}}{4\omega^2} \\ &= -\frac{m^2\ddot{a}}{4a\omega^3} + \frac{3\xi\ddot{a}}{a\omega} - \frac{\ddot{a}}{2a\omega} + \frac{5m^4\dot{a}^2}{8a^2\omega^5} - \frac{m^2\dot{a}^2}{2a^2\omega^3} + \frac{3\xi\dot{a}^2}{a^2\omega} - \frac{\dot{a}^2}{2a^2\omega} + \frac{s^2}{2\omega} \ , \end{aligned} \tag{3.367}$$

and at fourth adiabatic order, the result is

$$\begin{aligned} \omega^{(4)} &= \frac{2(s^2+\sigma)\omega\omega^{(2)} + 3/2(\dot{\omega}^{(2)}\dot{\omega}) - 1/2\left[\ddot{\omega}^{(2)}\omega + \ddot{\omega}\omega^{(2)}\right] - s\omega^2(\omega^{(2)})^2}{2\omega^3} \\ &= -\frac{1105\dot{a}^4m^8}{128a^4\omega^{11}} + \frac{221\ddot{a}\dot{a}^2m^6}{32a^3\omega^9} + \frac{221\dot{a}^4m^6}{16a^4\omega^9} - \frac{7\dot{a}a^{(3)}m^4}{8a^2\omega^7} - \frac{25s^2\dot{a}^2m^4}{16a^2\omega^7} - \frac{19\ddot{a}^2m^4}{32a^2\omega^7} - \frac{75\xi\dot{a}^2\ddot{a}m^4}{8a^3\omega^7} \\ &- \frac{111\dot{a}^2\ddot{a}m^4}{16a^3\omega^7} - \frac{75\xi\dot{a}^4m^4}{8a^4\omega^7} - \frac{69\dot{a}^4m^4}{16a^4\omega^7} + \frac{9\ddot{a}^2\xi m^2}{4a^2\omega^5} + \frac{15\dot{a}a^{(3)}\xi m^2}{4a^2\omega^5} + \frac{18\ddot{a}\dot{a}^2\xi m^2}{a^3\omega^5} + \frac{9\dot{a}^4\xi m^2}{2a^4\omega^5} \\ &+ \frac{5s\dot{a}\dot{s}m^2}{4a\omega^5} + \frac{3\ddot{a}s^2m^2}{8a\omega^5} + \frac{a^{(4)}m^2}{16a\omega^5} + \frac{2s^2\dot{a}^2m^2}{a^2\omega^5} + \frac{\ddot{a}^2m^2}{16a^2\omega^5} + \frac{\dot{a}a^{(3)}m^2}{16a^2\omega^5} - \frac{15\dot{a}^2\ddot{a}m^2}{16a^3\omega^5} - \frac{\dot{a}^4m^2}{4a^4\omega^5} \\ &+ \frac{3\ddot{a}\dot{a}^2\xi}{4a^3\omega^3} + \frac{3\dot{a}^4\xi}{2a^4\omega^3} + \frac{a^{(4)}}{8a\omega^3} - \frac{\dot{s}^2}{4\omega^3} - \frac{s\ddot{s}}{4\omega^3} - \frac{s^4}{8\omega^3} - \frac{3\xi s^2\ddot{a}}{2a\omega^3} - \frac{5s\dot{a}\dot{s}}{4a\omega^3} - \frac{3\xi a^{(4)}}{4a\omega^3} - \frac{3\xi s^2\dot{a}^2}{2a^2\omega^3} \\ &- \frac{9\xi^2\ddot{a}^2}{2a^2\omega^3} - \frac{s^2\dot{a}^2}{4a^2\omega^3} - \frac{3\xi\ddot{a}^2}{4a^2\omega^3} + \frac{\ddot{a}^2}{4a^2\omega^3} - \frac{15\xi\dot{a}a^{(3)}}{4a^2\omega^3} + \frac{5\dot{a}a^{(3)}}{8a^2\omega^3} - \frac{9\xi^2\dot{a}^2\ddot{a}}{a^3\omega^3} + \frac{\ddot{a}\dot{a}^2}{8a^3\omega^3} \\ &- \frac{9\xi^2\dot{a}^4}{2a^4\omega^3} - \frac{\dot{a}^4}{8a^4\omega^3} \ . \end{aligned} \tag{3.368}$$

Expressions for the subtraction terms in conformal time have been obtained in [137]. Here we will briefly sketch the renormalization counterterms associated to the UV divergences of the stress-energy tensor and the variance $\langle\phi^2\rangle$. We follow a strategy similar to the one used in Section 3.8. The Lagrangian density with the required renormalization counterterms is

$$\begin{aligned}\mathscr{L} &= \mathscr{L}_m + \sqrt{-g}\left[\frac{1}{2}g^{\mu\nu}\nabla_\mu\Phi\nabla_\nu\Phi - \frac{m^2}{2}\Phi^2 - \frac{\lambda}{4!}\Phi^4 - \frac{1}{2}\xi_2 R\Phi^2 - \frac{1}{8\pi G}\Lambda + \frac{1}{16\pi G}R + \alpha R^2\right] \\ &+ \sqrt{-g}\left[\frac{1}{2}\delta Z g^{\mu\nu}\nabla_\mu\Phi\nabla_\nu\Phi - \frac{\delta m^2}{2}\Phi^2 - \frac{\delta\lambda}{4!}\Phi^4 - \frac{1}{2}\delta\xi_2 R\Phi^2 - \frac{1}{8\pi}\delta\Lambda + \frac{1}{16\pi}\delta G^{-1}R + \delta\alpha R^2\right]\end{aligned} \tag{3.369}$$

One can check that the above counterterms are enough to absorb all the UV divergences that emerge in the quantization of the scalar field. We note that, due to the symmetry $\Phi \to -\Phi$ of the matter Lagrangian, counterterms of the form $R\Phi$, Φ, Φ^3 are absent. However, a higher-derivative term of the form R^2 is now necessary, which did not appear for the Dirac field in a FLRW spacetime.

We assume the conformal coupling to the curvature $\xi = 1/6$ in order to make the classical theory conformal invariant. For a massive field we have

$$g^{\mu\nu}T_{\mu\nu} - h^2\Phi^2\phi^2 = m^2\phi^2 \ . \tag{3.370}$$

Classical conformal invariance is obtained then when $m^2 = 0$. In adiabatic regularization the conformal anomaly is computed by taking the massless limit

$$C_s = g^{\mu\nu}\langle T_{\mu\nu}\rangle - h^2\Phi^2\langle\phi^2\rangle = -\lim_{m^2\to 0} m^2\langle\phi^2\rangle^{(4)} = -\lim_{m^2\to 0} m^2(4\pi a^3)^{-1}\int_0^\infty dk k^2 (W_k^{-1}(t))^{(4)} \ ,$$

where $(W_k^{-1}(t))^{(4)} = \omega^{-3}(\omega^{(2)})^2 - \omega^{-2}\omega^{(4)}$ is the fourth order term in the adiabatic expansion of W_k^{-1}. Note that here, $\langle\phi^2\rangle^{(4)}$ is evaluated including fourth-order adiabatic subtractions. This is different to the physical vacuum expectation value $\langle\phi^2\rangle_{ren}$, which has to be evaluated with subtractions only up to second order. This is why only the purely fourth order adiabatic piece contributes to the anomaly. The explicit expression of $(W_k^{-1})^{(4)}$ for arbitrary ξ is

$$\begin{aligned}(W_k^{-1})^{(4)}(t) = &+ \frac{1155\dot{a}^4m^8}{128a^4\omega^{13}} - \frac{231\dot{a}^2\ddot{a}m^6}{32a^3\omega^{11}} - \frac{231\dot{a}^4m^6}{16a^4\omega^{11}} + \frac{105\xi\dot{a}^4m^4}{8a^4\omega^9} + \frac{63\dot{a}^4m^4}{16a^4\omega^9} + \frac{35s^2\dot{a}^2m^4}{16a^2\omega^9} \\ &+ \frac{105\ddot{a}\xi\dot{a}^2m^4}{8a^3\omega^9} + \frac{105\ddot{a}\dot{a}^2m^4}{16a^3\omega^9} + \frac{7a^{(3)}\dot{a}m^4}{8a^2\omega^9} + \frac{21\ddot{a}^2m^4}{32a^2\omega^9} + \frac{3\dot{a}^4m^2}{4a^4\omega^7} + \frac{27\ddot{a}\dot{a}^2m^2}{16a^3\omega^7} \\ &- \frac{5\dot{a}s\dot{s}m^2}{4a\omega^7} - \frac{5s^2\ddot{a}m^2}{8a\omega^7} - \frac{a^{(4)}m^2}{16a\omega^7} - \frac{5\dot{a}^2s^2m^2}{2a^2\omega^7} - \frac{15\xi\ddot{a}^2m^2}{4a^2\omega^7} - \frac{15\dot{a}\xi a^{(3)}m^2}{4a^2\omega^7} + \frac{3\ddot{a}^2m^2}{16a^2\omega^7} \\ &- \frac{\dot{a}a^{(3)}m^2}{16a^2\omega^7} - \frac{45\dot{a}^2\xi\ddot{a}m^2}{2a^3\omega^7} - \frac{15\dot{a}^4\xi m^2}{2a^4\omega^7} + \frac{27\xi^2\dot{a}^4}{2a^4\omega^5} + \frac{3\dot{a}^4}{8a^4\omega^5} + \frac{9s^2\xi\dot{a}^2}{2a^2\omega^5} + \frac{27\ddot{a}\xi^2\dot{a}^2}{a^3\omega^5} \\ &+ \frac{3\ddot{a}\dot{a}^2}{8a^3\omega^5} + \frac{\dot{s}^2}{4\omega^5} + \frac{15a^{(3)}\xi\dot{a}}{4a^2\omega^5} + \frac{5s\dot{a}\dot{s}}{4a\omega^5} + \frac{s\ddot{s}}{4\omega^5} + \frac{3s^4}{8\omega^5} - \frac{s^2\ddot{a}}{2a\omega^5} + \frac{9s^2\ddot{a}\xi}{2a\omega^5} + \frac{3a^{(4)}\xi}{4a\omega^5} \\ &- \frac{a^{(4)}}{8a\omega^5} + \frac{27\ddot{a}^2\xi^2}{2a^2\omega^5} - \frac{\dot{a}^2s^2}{4a^2\omega^5} - \frac{9\xi\ddot{a}^2}{4a^2\omega^5} - \frac{5\dot{a}a^{(3)}}{8a^2\omega^5} - \frac{27\dot{a}^2\xi\ddot{a}}{4a^3\omega^5} - \frac{9\dot{a}^4\xi}{2a^4\omega^5} \ .\end{aligned} \tag{3.371}$$

The integral in comoving momenta is finite and independent of the mass. Assuming now $\xi = 1/6$, the result is

$$C_s = \frac{a^{(4)}}{480\pi^2 a} + \frac{\ddot{a}^2}{480\pi^2 a^2} - \frac{s\dot{a}\dot{s}}{16\pi^2 a} + \frac{a^{(3)}\dot{a}}{160\pi^2 a^2} - \frac{\dot{a}^2\ddot{a}}{160\pi^2 a^3} - \frac{s\ddot{s}}{48\pi^2} - \frac{\dot{s}^2}{48\pi^2} - \frac{s^4}{32\pi^2} \ . \tag{3.372}$$

We can rewrite the expression in terms of covariant scalar terms as

$$C_s = \frac{1}{2880\pi^2}\left\{\Box R - \left(R^{\mu\nu}R_{\mu\nu} - \frac{1}{3}R^2\right)\right\} - \frac{h^2}{48\pi^2}(\Phi\Box\Phi + \nabla^\mu\Phi\nabla_\mu\Phi + \frac{3h^2}{2}\Phi^4) \ , \tag{3.373}$$

which is the result given in equation (3.205).

G. Adiabatic expansion for Klen-Gordon fields. Master formula.

In this section we show the generic expression for the nth contribution in the WKB adiabatic expansion given by (3.222). Introducing the ansatz (3.221) into the equation of motion for the modes, one finds the following equation [48, 157],

$$W_k^2 = \omega^2 + \sigma + \frac{3}{4}\frac{\dot{W_k}^2}{W_k^2} - \frac{1}{2}\frac{\ddot{W}_k}{W_k} \ , \tag{3.374}$$

where

$$\sigma = \left(6\xi - \frac{3}{4}\right)\left(\frac{\dot{a}}{a}\right)^2 + \left(6\xi - \frac{3}{2}\right)\frac{\ddot{a}}{a} \ . \tag{3.375}$$

Equation (3.374) can be solved *algebraically* by iteration for initial value $\omega^{(0)} \equiv \omega = \sqrt{(k/a)^2 + m^2}$. Performing the calculation up to nth adiabatic order it can be shown that

$$\begin{aligned}
\omega^{(n)} \ = \ & \frac{1}{2\omega^3}\Bigg\{\omega^2\left[(\omega^{(n/2)})^2 + 2\sum_{i=2}^{n/2-1}\omega^{(i)}\omega^{(n-i)}\right] + \sigma\left[(\omega^{(n/2-1)})^2 + 2\sum_{i=0}^{n/2-2}\omega^{(i)}\omega^{(n-2-i)}\right] \\
& + \frac{3}{4}\left[(\dot{\omega}^{(n/2-1)})^2 + 2\sum_{i=0}^{n/2-2}\dot{\omega}^{(i)}\dot{\omega}^{(n-2-i)}\right] - \frac{1}{2}\left[\ddot{\omega}^{(n/2-1)}\omega^{(n/2-1)} + \sum_{i=0}^{n/2-2}\left(\ddot{\omega}^{(i)}\omega^{(n-2-i)} + \omega^{(i)}\ddot{\omega}^{(n-2-i)}\right)\right] \\
& - \Bigg[6\sum_{i=0}^{n/4-1}(\omega^{(i)})^2(\omega^{(n/2-i)})^2 + 4\sum_{k=0}^{n/2-1}(\omega^{(k)})^2\sum_{i=2}^{n/2-k-1}\omega^{(i)}\omega^{(n-i-2k)} + 4\sum_{k=2}^{n/2-1}(\omega^{(k)})^2\sum_{i=0}^{n/2-k-1}\omega^{(i)}\omega^{(n-i-2k)} \\
& + 8\sum_{i=0}^{n/4-2}\omega^{(i)}\sum_{j=i+2}^{n/2-2-i}\omega^{(j)}\sum_{k=j+2}^{n-j-2i-2}\omega^{(k)}\omega^{(n-k-i-j)} + 8\sum_{k=0}^{n/4-3/2}(\omega^{(k)})^2\sum_{i=k+2}^{n/2-k-1}\omega^{(i)}\omega^{(n-i-2k)} + (\omega^{(n/4)})^4\Bigg]
\end{aligned} \tag{3.376}$$

with $\omega^{(s)} = 0$ for $s < 0$ or s being a fractional number. With this formula we can recover $\omega^{(s)} = 0$, for s being an odd integer, and the corresponding expressions for orders 2 and 4 from [48, 157],

$$\omega^{(2)} \ = \ \frac{1}{2}\omega^{-1/2}\frac{d^2}{dt^2}\omega^{-1/2} + \frac{1}{2}\omega^{-1}\sigma \ , \tag{3.377}$$

$$\omega^{(4)} \ = \ \frac{1}{4}\omega^{(2)}\omega^{-3/2}\frac{d^2}{dt^2}\omega^{-1/2} - \frac{1}{2}\omega^{-1}(\omega^{(2)})^2 - \frac{1}{4}\omega^{-1/2}\frac{d^2}{dt^2}\left[\omega^{-3/2}\omega^{(2)}\right] \ , \tag{3.378}$$

as well. In general, (3.376) allows us to obtain any $\omega^{(n)}$ in terms of lower order adiabatic terms and its derivatives.

H. Adiabatic expansion for Dirac fields. Master formula.

In this section we present the generic expressions for the nth contribution in the Dirac adiabatic expansion given by (3.257)-(3.258). Introducing these expressions into the equation of motion for the modes, (3.9), one gets a set of coupled *algebraic* equations (3.19):

$$(\omega - m)G = (\Omega-\omega)F + i\dot{F} - \frac{im\dot{\omega}}{2\omega(\omega+m)}F + (\omega-m)F\ , \tag{3.379}$$

$$(\omega + m)F = (\Omega-\omega)G + i\dot{G} + \frac{im\dot{\omega}}{2\omega(\omega-m)}G + (\omega+m)G\ , \tag{3.380}$$

$$2\omega = (\omega+m)FF^* + (\omega-m)GG^*\ , \tag{3.381}$$

which can be solved *algebraically* by iteration for initial values $F^{(0)} = G^{(0)} = 1$ and $\omega^{(0)} = \omega$. The general algorithm to compute the three fundamental objects [notice that $G(-m)$ satisfies the same equations as $F(m)$, so we take $G(-m) = F(m)$] is provided by

$$\begin{aligned} \omega^{(n)} &= -\frac{m}{\omega}\left\{\sum_{l=1}^{n-1}\omega^{(l)}F^{(n-l)} + i\dot{F}^{(n-1)} - \frac{im\dot{\omega}}{2\omega(\omega+m)}F^{(n-1)}\right\} \\ &+ \left(1-\frac{m}{\omega}\right)\left\{-\frac{i}{2}\left[\dot{F}^{(n-1)}+\dot{G}^{(n-1)}\right] - \frac{1}{2}\sum_{l=1}^{n-1}\omega^{(l)}\left[F^{(n-l)}+G^{(n-l)}\right] + \frac{im\dot{\omega}}{4\omega}\left[\frac{F^{(n-1)}}{(\omega+m)} - \frac{G^{(n-1)}}{(\omega-m)}\right]\right\}\ , \end{aligned} \tag{3.382}$$

$$\begin{aligned} ReF^{(n)}(m) &= \frac{\delta_{n0}}{2} - \frac{1}{4\omega}\sum_{l=1}^{n-1}\left[F^{(l)}F^{*(n-l)}(\omega+m) + G^{(l)}G^{*(n-l)}(\omega-m)\right] + \frac{1}{2\omega}Im\dot{F}^{(n-1)}(m) \\ &-\frac{1}{2\omega}\sum_{l=1}^{n}\omega^{(l)}ReF^{(n-l)}(m) - \frac{m\dot{\omega}}{4\omega^2(m+\omega)}ImF^{(n-1)}(m)\ , \end{aligned} \tag{3.383}$$

$$ImF^{(n)}(m) = ImG^{(n)}(m) - \frac{1}{\omega-m}\left\{\sum_{l=1}^{n}\omega^{(l)}ImF^{(n-l)}(m) + Re\dot{F}^{(n-1)}(m) - \frac{m\dot{\omega}}{2\omega(\omega+m)}ReF^{(n-1)}(m)\right\} \tag{3.384}$$

with $F = ReF(m) + iImF(m)$ and $G = ReG(m) + iImG(m)$. Notice that there is an inherent ambiguity in the formalism reflected in the choice for $ImG(m)$, but it can be explicitly seen that it does not affect the observables such as $\langle\bar{\psi}\psi\rangle$ or $\langle T_{\mu\nu}\rangle$ [**?**]. The simplest way to remove the ambiguities is to assume $ImG^{(n)}(m) = -ImF^{(n)}(m)$. Detailed expressions for the first adiabatic contributions can be found in [126, 127]. In general, (3.382)-(3.384) allow us to obtain any Dirac adiabatic contribution in terms of lower order adiabatic terms and its derivatives.

I. DeWitt-Schwinger a_4 coefficient.

We give here the result for the a_4 DeWitt coefficient for a spatially flat FLRW spacetime obtained with the adiabatic regularization method (3.376). It is used to illustrate the efficiency of the adiabatic formalism to get the higher-order DeWitt coefficients, as discused in 3.12.

$$
\begin{aligned}
a_4(x) \;=\; & \frac{29\dot a^8}{120a^8}-\frac{379\dot a^6\ddot a}{210a^7}+\frac{899\dot a^4\ddot a^2}{280a^6}+\frac{83\dot a^2\ddot a^3}{35a^5}-\frac{13\ddot a^4}{21a^4}+\frac{47\dot a^5\dddot a}{70a^6}+\frac{2\dot a^3\ddot a\dddot a}{3a^5}-\frac{103\dot a\ddot a^2\dddot a}{28a^4}-\frac{647\dot a^2\dddot a^2}{840a^4}\\
&+\frac{103\ddot a^2\ddddot a}{210a^3}-\frac{2\dot a^4\ddddot a}{21a^5}-\frac{93\dot a^2\ddot a\ddddot a}{70a^4}+\frac{34\ddot a^2\ddddot a}{105a^3}+\frac{199\dot a\dddot a\ddddot a}{420a^3}+\frac{11\ddddot a^2}{504a^2}-\frac{13\dot a^3a^{(5)}}{210a^4}+\frac{41\dot a\ddot a a^{(5)}}{140a^3}\\
&+\frac{29\dddot a a^{(5)}}{1260a^2}+\frac{3\dot a^2a^{(6)}}{70a^3}+\frac{13\ddot a a^{(6)}}{1260a^2}-\frac{\dot a a^{(7)}}{126a^2}-\frac{a^{(8)}}{630a}-\frac{7\xi\dot a^8}{5a^8}-\frac{39\xi^2\dot a^8}{5a^8}+\frac{36\xi^3\dot a^8}{a^8}+\frac{54\xi^4\dot a^8}{a^8}\\
&+\frac{383\dot a^6\ddot a}{20a^7}-\frac{15\xi^2\dot a^6\ddot a}{a^7}-\frac{234\xi3\dot a^6\ddot a}{a^7}+\frac{216\xi^4\dot a^6\ddot a}{a^7}-\frac{8123\xi\dot a^4\ddot a^2}{140a^6}+\frac{2859\xi^2\dot a^4\ddot a^2}{10a^6}\\
&-\frac{432\xi^3\dot a^4\ddot a^2}{a^6}+\frac{324\xi^4\dot a^4\ddot a^2}{a^6}-\frac{254\xi\dot a^2\ddot a^3}{15a^5}+\frac{264\xi^2\dot a^2\ddot a^3}{5a^5}-\frac{180\xi^3\dot a^2\ddot a^3}{a^5}+\frac{216\xi^4\dot a^2\ddot a^3}{a^5}+\frac{523\xi\ddot a^4}{105a^4}\\
&-\frac{81\xi^2\ddot a^4}{10a^4}-\frac{18\xi^3\ddot a^4}{a^4}+\frac{54\xi^4\ddot a^4}{a^4}-\frac{211\xi\dot a^5\dddot a}{20a^6}+\frac{201\xi^2\dot a^5\dddot a}{5a^6}-\frac{18\xi^3\dot a^5\dddot a}{a^6}+\frac{53\xi\dot a^3\ddot a\dddot a}{7a^5}-\frac{69\xi^2\dot a^3\ddot a\dddot a}{a^5}\\
&+\frac{72\xi^3\dot a^3\ddot a\dddot a}{a^5}+\frac{439\xi\dot a\ddot a^2\dddot a}{14a^5}-\frac{84\xi^2\dot a\ddot a^2\dddot a}{a^4}+\frac{90\xi^3\dot a\ddot a^2\dddot a}{a^4}+\frac{6\xi\dot a^2\dddot a^2}{a^4}-\frac{147\xi^2\dot a^2\dddot a^2}{10a^4}\\
&+\frac{18\xi^3\dot a^2\dddot a^2}{a^4}-\frac{51\xi^3\ddot a\dddot a^2}{20a^3}-\frac{3\xi^2\ddot a\dddot a^2}{a^3}+\frac{18\xi^3\ddot a\dddot a^2}{a^3}+\frac{11\xi\dot a^4\ddddot a}{4a^5}-\frac{15\xi^2\dot a^4\ddddot a}{a^5}+\frac{18\xi^3\dot a^4\ddddot a}{a^5}\\
&+\frac{157\xi\dot a^2\ddot a\ddddot a}{14a^4}-\frac{153\xi^2\dot a^2\ddot a\ddddot a}{5a^4}+\frac{36\xi^3\dot a^2\ddot a\ddddot a}{a^4}-\frac{19\xi\ddot a^2\ddddot a}{14a^3}-\frac{24\xi^2\ddot a^2\ddddot a}{5a^3}+\frac{18\xi^3\ddot a^2\ddddot a}{a^3}\\
&-\frac{237\xi\dot a\dddot a\ddddot a}{70a^3}+\frac{27\xi^2\dot a\dddot a\ddddot a}{5a^3}-\frac{39\xi\ddddot a^2}{140a^2}+\frac{9\xi^2\ddddot a^2}{10a^2}+\frac{3\xi\dot a^3a^{(5)}}{10a^4}+\frac{41\dot a\ddot a a^{(5)}}{140a^3}-\frac{15\xi\dot a\ddot a a^{(5)}}{7a^3}\\
&+\frac{18\xi^2\dot a\ddot a a^{(5)}}{5a^3}-\frac{12\xi\dddot a a^{(5)}}{35a^2}+\frac{6\xi^2\dddot a a^{(5)}}{5a^2}+\frac{3\dot a^2a^{(6)}}{70a^3}-\frac{23\xi\dot a^2a^{(6)}}{70a^3}+\frac{3\xi^2\dot a^2a^{(6)}}{5a^3}-\frac{6\xi\ddot a a^{(6)}}{35a^2}+\frac{3\xi^2\ddot a a^{(6)}}{5a^2}\\
&+\frac{\xi\dot a a^{(7)}}{28a^2}+\frac{\xi a^{(8)}}{140a^2}\,. \qquad (3.385)
\end{aligned}
$$

$\delta\ \epsilon\ \lambda\ \varphi \rightarrow \mathscr{B}\ \mathscr{F}\ \mathscr{L}\ \mathscr{P}\ \zeta\ \mathscr{T}\ \mathscr{R}_c\ \mathscr{G}\ \mathscr{H}\ \nabla\ \mathbf{x}\ \mathbf{k}\ \mathbf{p}\ \widehat{\mathscr{R}}\ \mathscr{H}\ \mathscr{A}\ \mathscr{Q}\ \mathscr{O}\ \mathscr{M}\ \mathscr{N}\ \mathscr{S}\ \mathbf{p}\ \mathbf{k}$

Chapter 4

Loop corrections during inflation and bounds from CMB data

During single-field inflation, spectator or hidden fields – i.e. those that only propagate in the gravitational background and do not couple to the inflaton – can typically affect CMB observables only through quantum fluctuations. After renormalizing background quantities, such as slow-roll parameter ϵ or Planck mass M_{pl} (that are fixed by observations at some pivot scale), all that remains are logarithmic runnings that are suppressed by both M_{pl} and slow roll parameters. In this paper we show how a large number of spectator fields can overcome this suppression and induce an observable running in the tensor two point function. As a consequence, one can infer bounds on the hidden field content of the universe from bounds on the tensor tilt, assuming primordial tensors are ever detected. We point out that the bounds obtained from spectral running are more competitive than the naive bound inferred from requiring inflation to occur below the strong coupling scale of gravity if the eventual measurement of the tensor tilt is negative, and if we can bound deviations from the tensor to scalar consistency relation to within the percent level. We finally discuss some phenomenological scenarios where this idea could have potential implications, such as constructions that address the hierarchy problem in the standard model with a huge number of species.

The work presented in this chapter is done in collaboration with R. Durrer and S. P. Patil [63].

4.1 Motivation

Observations strongly suggest that our Universe went through an early phase of quasi-exponential expansion, that we call inflation. Such an inflationary stage not only solves the horizon and flatness puzzles [108, 129], it also originates in a natural way an almost scale invariant spectrum of matter density fluctuations [38, 109, 112, 139, 179] in agreement with measurements from the cosmic microwave background (CMB). These primordial inhomogeneities were generated as quantum vacuum fluctuations that were forced out of the horizon by the huge gravitational dynamics of the Universe and subsequently squeezed, resulting in their phase coherence. The inflationary expansion also amplifies vacuum fluctuations of the transverse traceless part of the metric, leading to the generation of primordial gravitons [180] as well as fluctuations of all other fields present in the quantum vacuum regardless they couple directly to the inflaton or not.

In this chapter we examine in detail the contribution of fields that one would naturally be tempted to neglect during inflation: hidden or spectator fields, defined as fields that only propagate in the curved spacetime background and have no direct couplings to the inflaton. Typically, such fields would only serve to renormalize background quantities[1] and produce unconceivable small (i.e. Planck suppressed) logarithmic dependence of momentum (runnings) in the cosmological two-point functions. Nevertheless, for sufficiently large amounts, their effects can add up to an observable running of the spectral index of the two point function of the tensor perturbations, consistently inferable via a "large N" expansion that allows us to resum a restricted class of diagrams. As we shall argue, the analogue running produced for correlators of curvature perturbations remains indistinguishable from the other expected contributions, though, since the relative suppression by factors of ϵ is too great to be overcome by large N and still consistent with being below the strong coupling scale of gravity.

Taking into account this observation, we can use it to infer bounds on the possible number of hidden fields present in the universe with masses below the scale of inflation through bounds on the tensor to scalar consistency relation, provided primordial tensors are ever observed [2] For simplicity, we focus on hidden scalars, although the argument follows straightforwardly to particles of any spin. We find that any upper bound on the

[1] Whose effects therefore would simply be absorbed into physical measurements of quantities such as $\epsilon := -\dot{H}/H^2$ (e.g. through the detection of primordial tensors) and its derivatives or the ratio H^2/M_{pl}^2, all of which denote renormalized quantities.

[2] Although fields with masses much greater than the Hubble scale during inflation also contribute to the running their spectrum is very suppressed at long wavelengths and so will not contribute to the bounds derived here.

tensor to scalar consistency relation

$$n_T + \frac{r_\star}{8} \leq \xi \tag{4.1}$$

for some ξ, translates into a bound on the number of hidden species as

$$N \leq \frac{\xi}{r_\star^2} \Delta_\zeta^{-1} \tag{4.2}$$

where $\Delta_\zeta \approx 2.44 \times 10^{-9}$ [166] is the amplitude of the spectrum of curvature perturbations at the pivot scale where we determine the tensor to salar ratio. Assuming the value of $r_\star$ to be that of his current experimental bound, $r \sim \mathcal{O}(10^{-1})$, the best we can bound N is by

$$N \lesssim \frac{10^9}{r_\star^2} \xi \sim 10^{11} \times \xi \tag{4.3}$$

This result will be only interesting if it gives a stronger bound than the one arising from the fact that inflation must have occurred below the scale at which gravity becomes strongly coupled: $H^2 \lesssim M_{pl}^2/N$, [79]. We shall argue that, although in order to get a better bound it is necessary that $n_T < 0$ in future measurements, the bigger the eventual measurement of $|n_T|$ is, the stronger bound we get (as compared to the strong coupling). As we shall discuss, one can look to bound the parameter space of a variety of models that attempt to address the hierarchy problem with a large number of sectors.

Notation. In what follows, we shall use the $(+,+,+)$ convention of [136]. We shall consider a flat FLRW metric in cartesian coordinates

$$ds^2 = a^2(\eta)\left[-d\eta^2 + \delta_{ij} dx^i dx^j\right] = g_{\mu\nu} dx^\mu dx^\nu \tag{4.4}$$

where η denotes conformal time, while the physical time t is given by $dt = a d\eta$. Derivatives with respect to η are denoted by a prime and those with respecto to t by an overdot. The physical Hubble parameter is $H = \frac{\dot{a}}{a}$ while the conformal one is $\mathcal{H} = \frac{a'}{a} = \dot{a}$.

4.2 Outline of the calculation

In this section we present an outline of our calculations and their results with the details deferred to several appendices. We consider an inflationary Universe with an inflaton ϕ taken to be the only field with an evolving background (hence energy density) and N additional hidden scalar fields χ_n minimally coupled to gravity and taken to be in their respective adiabatic vacuum states. We only consider hidden fields with mass $m^2 \ll H^2$, which can therefore be treated as effectively massless but are quantum mechanically excited during inflation. By assumption the χ_n have no non-gravitational interactions. The action is then given by:

$$S = \frac{M_{pl}^2}{2}\int d^4x\sqrt{-g}R[g] - \frac{1}{2}\int d^4x\sqrt{-g}\left[\partial_\mu\phi\partial^\mu\phi + 2V(\phi) + \sum_{n=1}^{N}\left(\partial_\mu\chi_n\partial^\mu\chi_n + m_n^2\chi_n^2\right)\right] \quad (4.5)$$

where $M_{pl} = (8\pi G)^{-1/2}$ is the reduced Planck mass[3]. We presume the background to be quasi de-Sitter, such that

$$\epsilon := \frac{\dot{\phi}_0^2}{2H^2M_{pl}^2} = -\frac{\dot{H}}{H^2} \ll 1, \qquad \delta := \frac{\ddot{\phi}_0}{\dot{\phi}_0 H} = \frac{\ddot{H}}{2\dot{H}H} \ll 1, \quad (4.6)$$

so that $H^2 = V(\phi_0)/(3M_{pl}^2) \sim$ const, and for completeness we introduce the higher order slow roll parameters ϵ_i, defined by

$$\epsilon_1 \equiv \epsilon, \qquad \epsilon_{i+1} = \frac{\dot{\epsilon}_i}{H\epsilon_i}, \quad i \geq 1. \quad (4.7)$$

Note that $\epsilon_2 = 2(\delta+\epsilon)$. These slow roll parameters are then of the same order of magnitude $\epsilon_{i+1} \sim \epsilon_i \sim \epsilon$.

In order to discuss perturbations around this background, we first ADM decompose the metric as

$$ds^2 = -N^2dt^2 + h_{ij}(dx^i + N^i dt)(dx^j + N^j dt), \quad (4.8)$$

and work in comoving gauge, defined to be the foliation in which we have gauged away the inflaton fluctuations. In this gauge, the only dynamical degrees of freedom are contained in the 3-metric h_{ij} which has now acquired, or 'eaten' the scalar polarization that was the inflaton fluctuation [57]

$$\phi(t,x) = \phi_0(t), \quad (4.9)$$

$$h_{ij}(t,x) = a^2(t)e^{2\zeta(t,x)}\hat{h}_{ij}, \qquad \hat{h}_{ij} = \exp\left[\gamma_{ij}\right] \quad (4.10)$$

where $\gamma^i_i = \partial_i\gamma^i_j = 0$ is the (transverse traceless) graviton, and ζ is the comoving curvature perturbation. The quasi dS background then results in a nearly scale invariant spectrum of curvature perturbations [77, 138, 189]

$$P_\zeta(k) = \frac{H^2}{8\pi^2\epsilon M_{pl}^2}\left(\frac{k}{k_\star}\right)^{n_s-1}, \qquad n_s - 1 = -4\epsilon - 2\delta = -2\epsilon - \epsilon_2. \quad (4.11)$$

[3]This so far bare quantity will also end up being renormalized via diagrams involving external graviton legs with loops of hidden fields. However, we only end up seeing this quantity through the dimensionless ratio H^2/M_{pl}^2 which we presume to be fixed through given any observation of primordial tensors at some fixed scale

In addition helicity 2 tensor perturbations of the metric are amplified from their initial quantum vacuum state leading to a power spectrum for primordial gravitational waves given by

$$P_\gamma(k) = \frac{2H^2}{\pi^2 M_{pl}^2}\left(\frac{k}{k_\star}\right)^{n_T}, \qquad n_T = -2\epsilon\,. \tag{4.12}$$

The ratio of these two quantities

$$r = \frac{P_\gamma(k_\star)}{P_\zeta(k_\star)} = 16\epsilon \tag{4.13}$$

defines the tensor to scalar ratio. Note that its value depends on the pivot scale $k_\star$. At present, no tensor perturbations have been identified in the observed CMB anisotropies and an upper limit of $r < 0.11$ has been derived for $k_\star = 0.002$ Mpc^{-1} [9]. We remind the reader that there are higher order corrections to the tilt of the scalar and tensor spectra that come from the background dynamics alone (see (4.51)-(4.55) below). We are interested in additional corrections to these from virtual effects of the hidden fields χ_n.

Diagrammatic preliminaries in the 'in-in' formalism

By following a perturbative study in powers of ϵ of the action (4.5) in comoving gauge (4.9) results in the quadratic action

$$S_{2,\zeta} = M_{pl}^2 \int d^4x\, a^3\, \epsilon \left[\dot\zeta^2 - \frac{1}{a^2}(\partial\zeta)^2\right] \tag{4.14}$$

$$S_{2,\chi} = \frac{1}{2}\int d^4x\, a^3 \left[\dot\chi_n\dot\chi_n - \frac{1}{a^2}\partial_i\chi_n\partial_i\chi_n - m_n^2\chi_n^2\right] \tag{4.15}$$

and the cubic interaction vertex

$$S_{3,\zeta\chi} = \int d^4x\, a^3\epsilon\left[\frac{\zeta}{2}\left(\dot\chi_n\dot\chi^n + \frac{1}{a^2}\partial_i\chi_n\partial_i\chi_n + m_n^2\chi_n^2\right) - \dot\chi_n\partial_i\chi_n\partial_i\partial^{-2}\dot\zeta\right] \tag{4.16}$$

$$S_{3,\gamma\chi} = \frac{1}{2}\int d^4x\, a\gamma_{ij}\partial_i\chi_n\partial_j\chi_n \tag{4.17}$$

The form of (4.16) – in particular its ϵ suppression – is not immediately obvious from naively expanding the original action (4.5) having solved for the lapse and shift constraints, which would result in an expression that is nominally unsuppressed in ϵ (4.74). However, as shown in detail in Appendix A, similar to what occurs for the cubic and higher order self interactions for ζ [131], enough integrations by parts makes manifest the fact that the $\zeta\chi\chi$ cubic interactions are in fact suppressed by an overall factor of ϵ. Similarly, interactions that are higher order in ζ will be sequentially suppressed by additional powers of ϵ, consistent with its nature as an order parameter parametrizing the breaking of time translational invariance by slow roll [58].

The standard approach at this stage would be to work in the Schwinger-Keldysh, or in-in formalism to calculate the finite time correlation functions

$$k^3\langle\zeta_{\mathbf{k}}(\tau)\zeta_{\mathbf{p}}(\tau)\rangle := 2\pi^2\delta^2(\mathbf{k}+\mathbf{p})\mathscr{P}_\zeta(k), \tag{4.18}$$

$$k^3\langle\gamma^r_{ij,\mathbf{k}}(\tau)\gamma^r_{ij,\mathbf{p}}(\tau)\rangle := 2\pi^2\delta^{rs}\delta^2(\mathbf{k}+\mathbf{p})\mathscr{P}_\gamma(k), \tag{4.19}$$

These quantities are of the form $\langle\mathscr{O}(\tau)\rangle$, where the angled brackets denote expectation values with a given initial density matrix (which we take to correspond to the Bunch-Davies vacuum in this work) unitarily evolved forward in the interaction picture with the Dyson operator

$$U(\tau,-\infty) = T\exp\left(-i\int_{-\infty}^{\tau} H_I(\tau')d\tau'\right), \tag{4.20}$$

where T denotes time ordering and where H_I is the interaction Hamiltonian (which is equal to minus the interaction Lagrangian eq. (4.16) for the interactions in question [188]). Explicitly then, the operator expectation value is a shorthand for

$$\langle\mathscr{O}(\tau)\rangle = \langle 0_{in}|\left[T\exp\left(-i\int_{-\infty}^{\tau} H_I(\tau')d\tau'\right)\right]^\dagger \mathscr{O}_0(\tau)\left[T\exp\left(-i\int_{-\infty}^{\tau} H_I(\tau')d\tau'\right)\right]|0_{in}\rangle \tag{4.21}$$

Reading right to left, one evidently evolves the Bunch-Davies vacuum from the initial time $-\infty$ to τ, acts with the operator $\mathscr{O}(\tau)$ at time τ and then evolves back to $-\infty$ (see Fig. 4.1). It can be shown that the above is equivalent to the expression [188]

$$\langle\mathscr{O}(\tau)\rangle = \sum_{n=0}^{\infty} i^n\int_{-\infty}^{\tau} d\tau_n\int_{-\infty}^{\tau_n} d\tau_{n-1}...\int_{-\infty}^{\tau_2} d\tau_1\langle[H_I(\tau_1),[H_I(\tau_2),...[H_I(\tau_n),\mathscr{O}(\tau)]...]]\rangle \tag{4.22}$$

provided one is mindful of how one selects the correct initial interacting vacuum [11].

Although useful for practical purposes, such an expectation value does not lend itself to the usual diagrammatic expansion one avails of when dealing with S-matrix elements. In order to implement this one can equivalently consider expressions like (4.21) as the product of an arbitrary operator $\mathscr{O}(\tau)$ with the unitary operator:

$$\langle\mathscr{O}(\tau)\rangle = \langle 0_{in}|T_C\left[exp\left(-i\oint H_I(\tau')d\tau'\right)\mathscr{O}(\tau)\right]|0_{in}\rangle \tag{4.23}$$

with the contour going from $-\infty\to\tau$ and back again, and with T_C denoting contour ordering with fields living on the reversed contour being treated as independent fields for intermediate manipulations, only being set equal to the original fields at the end of the calculation. Due to its formal similarity with an S-matrix element, the former does

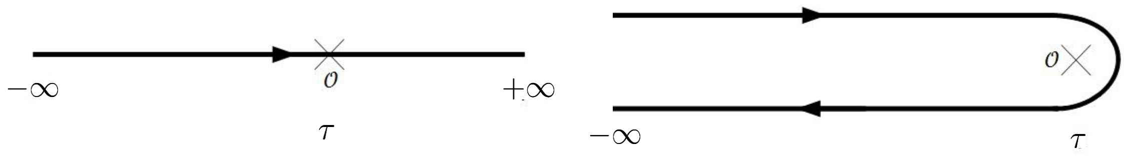

FIGURE 4.1. The S-matrix contour (left) compared to the Schwinger-Keldysh contour (right)

indeed lend itself to the usual diagrammatic expansion, which we will not make explicit use of in the following, but we find useful for building diagrammatic intuition.

Suppressing the difference between the fields that live on the future and past directed contours (as a result of which there are typically cancellations as one sums up relevant diagrams) as shorthand, one can nevertheless diagrammatically intuit the parametric and external momentum dependences of the various graphs that one can write down. At one loop, one has two possible contributions to the correction to the two point correlation function[4] as indicated in fig. 4.2. However only the diagram involving two cubic vertices results in any dependence on the external momenta[5] and hence contributions to the running of the spectral index, which is the object of our interest.

Some versions of this calculation in the in-in formalism already exist in the literature [188], though some conceptual mistakes were detected by [176]. We provide another version of the calculation done in the Heisenberg picture, after which we will turn to the main result of this paper – that it is possible to resum these diagrams in the large N limit so that we could consistently infer the running induced by a large number of hidden fields.

[4]There are also contributions from cubic interactions involving ζ alone, but these will be suppressed by two extra powers of ϵ [131, 177].

[5]The quartic seagull interactions contributes to wave-function renormalization which is accounted for in practice by fixing the (fully renormalized) expressions H^2/M_{pl}^2 and ϵ via the amplitude of the power spectrum and the tensor to scalar ratio at some pivot scale k_*.

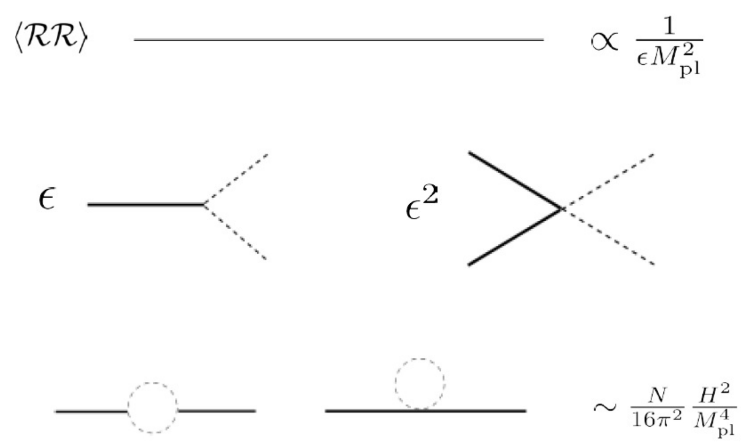

FIGURE 4.2. One loop corrections to $\langle\zeta\zeta\rangle$. Solid lines denote the curvature perturbation propagator, dashed lines denote the χ-propagator.

Heisenberg picture calculation

Let us consider the (classical) equation of motion for a given mode $\chi_{\mathbf{k}}$ of the Fourier mode $\mathbf{k}$ of a hidden scalar field in a FLRW universe (we neglect the mass term, $m \ll H$):

$$\ddot{\chi}_k + 3H\dot{\chi}_k + \frac{\mathbf{k}^2}{a^2}\chi_k = 0, \qquad \chi_k'' + 2Ha\chi_k' + k^2\chi_k = 0 \tag{4.24}$$

Assuming that the field be in the so-called Bunch-Davies vacuum initially (in the limit $\epsilon \to 0$ and $|k\eta| \gg 1$), (4.24) is solved by

$$\chi_k(t) = \frac{\sqrt{-\pi\eta}}{2a} H_\mu^{(1)}\left(-k\eta\right), \qquad \mu = \frac{3}{2} + \epsilon. \tag{4.25}$$

We assume each of these fields to contribute as a first-order perturbation in the perturbed Einstein's equations (see Appendix A). The $\mathbf{k}$-modes of the energy momentum tensor of a hidden field,

$$T^\chi_{\mu\nu} = -\frac{1}{2}\bar{g}_{\mu\nu}\partial_\sigma\chi\partial^\sigma\chi + \partial_\mu\chi\partial_\nu\chi\,,$$

can now be calculated as convolutions. They are second order on the fields χ, hence of second order in linear perturbation theory.

This second order energy momentum tensor also induces metric fluctuations, besides the fluctuations originated by the inflaton field itself. It provides scalar (helicity zero), vector (helicity 1) and tensor (helicity 2) fluctuations which decouple in first order. Here we only consider scalar and tensor perturbations since vector perturbations decay after inflation.

Scalar perturbations

Inserting it into the perturbed Einstein equations in unitary gauge and using energy momentum conservation of the scalar field χ we find (see Appendix B for a detailed derivation),

$$\zeta'' + 2\frac{z'}{z}\zeta' + k^2\zeta = \frac{1}{2M_{pl}^2}a^2\left\{\rho - p + \frac{2}{3}\Pi_s + \frac{q}{a}\frac{\epsilon'}{\epsilon}\right\}, \qquad z = \frac{a\phi_0'}{\mathcal{H}}. \tag{4.26}$$

We can obtain the above equation of motion also from the cubic interaction eq. (4.16) derived in Appendix A:

$$S_{3,\zeta\chi} = \int d^4x\, a^3\epsilon\left\{\frac{\zeta}{2}\left[\dot{\chi}\dot{\chi} + \frac{1}{a^2}\partial_i\chi\partial_i\chi\right] + \dot{\zeta}\partial^{-2}\partial_i(\dot{\chi}\partial_i\chi)\right\} \tag{4.27}$$

$$= \int d^4x\, a^3\epsilon\{\zeta\rho - q\dot{\zeta}\} = \int d^4x\, a^3\epsilon\,\zeta\left[\rho + \dot{q} + 3Hq + \frac{\dot{\epsilon}}{\epsilon}q\right] \tag{4.28}$$

Using momentum conservation, the source term becomes $\rho + \dot{q} + 3Hq + \frac{\dot{\epsilon}}{\epsilon}q = \rho - p + \frac{2}{3}\Pi_s + \frac{\epsilon'}{\epsilon}\frac{q}{a}$, so that the variation of $S_{3,\zeta\chi}$ wrt. ζ simply yields the right hand side of (4.26). As is well known, the quadratic action for ζ yields the left hand side.

Setting the right hand side of eq. (4.26) to zero, results in the Mukhanov-Sasaki equation. Its solution with initial conditions given by the Bunch-Davies vacuum leads to the spectrum (4.11). This is the homogeneous part of the solution. The source term on the right hand side is determined by the scalar fields χ_n. Their energy density and pressure are given by ρ and p, q is the potential of the energy flux and Π_s is the scalar anisotropic stress potential, see Appendix B and C for all details. As the source term is second order in χ, the inhomogeneous contribution to the spectrum is fourth order in χ_n and can be computed using Wick's theorem for the Gaussian (free) fields χ_n. This calculation is presented in detail in Appendix D with the result

$$P_{\zeta_i}(k) \approx \frac{H^4}{15\times 8\pi^2 M_{pl}^4}\left[c + \log\frac{H_k}{H_\star}\right]\left(\frac{k}{k_\star}\right)^{n_s-1}. \tag{4.29}$$

The parameters $n_s - 1 = -4\epsilon - 2\delta + O(\epsilon^2)$ is the standard scalar spectral index and c is a constant; $k_\star$ is the pivot scale, the Hubble parameter and the slow roll parameters have to be evaluated at $H_\star \equiv H(k_\star)$. We denoted H with a subindex k inside the logarithm, the reason will be made clear in a couple of subsections. The index 'i' indicates that this is the contribution from the inhomogeneous solution sourced by the scale field χ_n. A total nomber of N such scalar fields contributes

$$NP_{\zeta_i}(k) \approx \frac{N\epsilon H^2}{15\pi^2 M_{pl}^2}\left[c + \log\frac{H_k}{H_\star}\right]P_\zeta(k). \tag{4.30}$$

As we show below, the first term $\propto c$ actually just contributes to a renormalisation of the slow roll parameter ϵ, it is not directly observable. However the second term induces a running of the spectral index which we shall discuss in a couple of subsection.

Tensor perturbations

Tensor perturbations are sourced by the tensor (i.e. transverse traceless) part of the anisotropic stress of $T_\chi^{\mu\nu}$. Denoting the tensor metric perturbations in Fourier space by $\gamma_{ij}(t,\mathbf{k})$, the perturbed Einstein equations are

$$\hat{\gamma}''_{ij} + 2\mathscr{H}\hat{\gamma}'_{ij} + k^2\hat{\gamma}_{ij} = \frac{2}{M_{pl}^2}\Pi^T_{ij}, \tag{4.31}$$

where Π^T_{ij} are the tensor modes of the anisotropic stress. Alternatively, one can identify the relevant cubic interaction term as in (4.83). In appendix E we compute the corresponding power spectrum with the result

$$NP_{\gamma_i}(k) \quad \approx \quad \frac{NH^4}{15\pi^2 M_{Pl}^4}\left[c + \log\frac{H}{H_\star}\right]\left(\frac{k}{k_\star}\right)^{n_T}, \tag{4.32}$$

where N denotes the number of hidden scalar fields χ present in the theory. Since no tensor perturbation have been detected so far, we can presently only derive an upper limit on the amplitude.

Running of the one-loop corrections

We notice that the original calculation of [188] and subsequent derivations in [11] using the in-in formalism obtained a $\log(k/H_\star)$ dependence instead of the $\log(H_k/H_\star)$ presented in formulas above. They concluded this way that one-loop corrections certainly contribute to the running of the power spectra. However, as realized in [176], the $\log(k/H_\star)$ dependence is only one of two contributions to the loop correction. An additional dependence $\log(H/k)$ also arises from corrections to the mode functions proportional to δ in the dimensional regularization process when one is dealing with $3+\delta$ spatial dimensions. This results in a $\log(k/H_\star) + \log(H/k) = \log(H/H_\star)$ dependence (see section 3.2 in ([176]) for a detailed discussion). Thus, after fixing renormalized quantities at some pivot scale $H_\star$, the dependence in (4.30) and (4.32) arises. Nevertheless, it is important to realize that $H \equiv H_k$ itself runs as a function of k as inflation progresses. As we shall see shortly the dependence is $\log k$ again, but suppressed with an extra factor of ϵ and an additional minus sign. So at the end the original conclusion that one-loop corrections to two-point functions induce a running in their spectra is still valid, though further suppressed by extra factors of slow-roll parameters.

The number of e-folds produced during the expansion at time t is defined to be $\mathcal{N}(t) := \log a(t)$. At the time t_k when the mode k crosses the horizon (defined by $k =: a(t_k)H_k$), the number of e-foldings is characterized by $\mathcal{N}_k \equiv \mathcal{N}(t_k) = \log k - \log H_k$. The evolution of the inflationary expansion can thus be parametrized by the modes k as well. On the other hand, recall that by definition ϵ can be written as $\frac{dH}{d\mathcal{N}} = -\epsilon H$. This can be integrated:

$$\log \frac{H_k}{H_\star} \sim -\int_{\mathcal{N}_*}^{\mathcal{N}_k} \epsilon(\mathcal{N}')d\mathcal{N}' \tag{4.33}$$

We can further obtain:

$$\log \frac{H_k}{H_\star} = \log \frac{k}{k_\star} - (\mathcal{N}_k - \mathcal{N}_\star) \tag{4.34}$$

Taking derivatives at both sides

$$\frac{\partial \mathcal{N}_k}{\partial \log k/k_\star} = 1 - \frac{\partial \log H_k/H_\star}{\partial \log k/k_\star} \tag{4.35}$$

and using the chain rule

$$\frac{\partial \log H_k/H_\star}{\partial \log k/k_\star} = \frac{\partial \mathcal{N}_k}{\partial \log k/k_\star}\frac{\partial \log H_k/H_\star}{\partial \mathcal{N}_k} = -\epsilon_k \frac{\partial \mathcal{N}_k}{\partial \log k/k_\star} = -\epsilon_k \left[1 - \frac{\partial \log H_k/H_\star}{\partial \log k/k_\star}\right] \tag{4.36}$$

Note so far this relation is exact. Therefore:

$$\frac{\partial \log H/H_k}{\partial k/k_\star} = -\frac{\epsilon_k}{1-\epsilon_k} \tag{4.37}$$

In the limit in which ϵ_k is small and constant we get

$$\log \frac{H_k}{H_\star} = -\epsilon \log \frac{k}{k_\star} + O(\epsilon^2) \tag{4.38}$$

This is the reason for which we denoted H a subindex k inside the logarithm. This contribution will be taken into account in the following discussions.

4.3 Observational limits on the number of hidden fields N

Running of the spectral index

As alluded to in the outline, in order to consider the contribution of hidden fields to the running of the spectral index, we only need to consider the contribution of the one loop diagram involving two cubic vertices as illustrated in Fig. 4.2. Nominally, each curvature perturbation propagator contributes a factor $1/(\epsilon M_{pl}^2)$ due to the non-canonically normalized nature of its kinetic term (4.14) whereas each $\zeta\chi\chi$ vertex is first order in ϵ due to the factor ϵ in the cubic action (4.16), while the $\zeta\zeta\chi\chi$ vertex will be second order in ϵ while each loop contributes a factor $\frac{NH^2}{16\pi^2}$ after being dimensionally regularized. Hence the 'seagull' contribution is of the same order in ϵ as the correction obtained from the cubic interactions – the extra factor of ϵ is compensated for by the fact that one fewer vertex is required. However, the loop momentum going around the seagull graph is independent of the incoming momenta and so cannot affect the dependence of the correlator of the external momentum and thus only contributes to wave-function renormalization – in practice accounted for by measurements of H^2/M_{pl}^2 and ϵ at some pivot scale k_*, the latter requiring an observation of primordial tensors.

Resummation in the large N limit

The correction to the power spectrum of the curvature perturbation at one loop from N hidden scalar fields is calculated in the appendices and shown in (4.30). The total power spectrum is given by

$$\begin{aligned} P_\zeta^{1-loop} &= P_\zeta + NP_{\zeta_i} = \frac{H^2}{8\pi^2\epsilon M_{pl}^2}\left(\frac{k}{k_\star}\right)^{n_s-1}\left(1-N\frac{\epsilon^2 H^2}{15M_{Pl}^2}\log\frac{k}{k_\star}\right) \\ &= \Delta\left(\frac{k}{k_\star}\right)^{n_s-1}\left[1-N\frac{\epsilon^2H^2}{15M_{pl}^2}\log\frac{k}{k_\star}\right], \end{aligned} \tag{4.39}$$

where $P_\zeta = \Delta\left(\frac{k}{k_\star}\right)^{n_s-1}$, is the standard measured power spectrum with amplitude $\Delta \equiv \frac{H^2}{8\pi^2\epsilon M_{pl}^2} \simeq 2.2\times 10^{-9}$ [166]. All quantities appearing above are presumed to have also taken wave function renormalization into account, with Δ and ϵ to be independently fixed by measurements of the amplitude of the power spectrum and the tensor to scalar ratio at k_*.

As illustrated in fig. 4.3 we see that among the two loop graphs indicated, the first one dominates in the limit of a large number of hidden fields due to its additional factor N. Hence, in the large N limit, one can perform a resummation of the basic one loop correction (4.39) such that we can consistently extract the running induced even in the limit where the one loop correction becomes comparable to the tree-level result:

$$P_{\zeta,tot} = \frac{\Delta\left(\frac{k}{k_*}\right)^{n_s-1}}{1+N\frac{\epsilon^2H^2}{15M_{pl}^2}\log\frac{k}{k_\star}}. \tag{4.40}$$

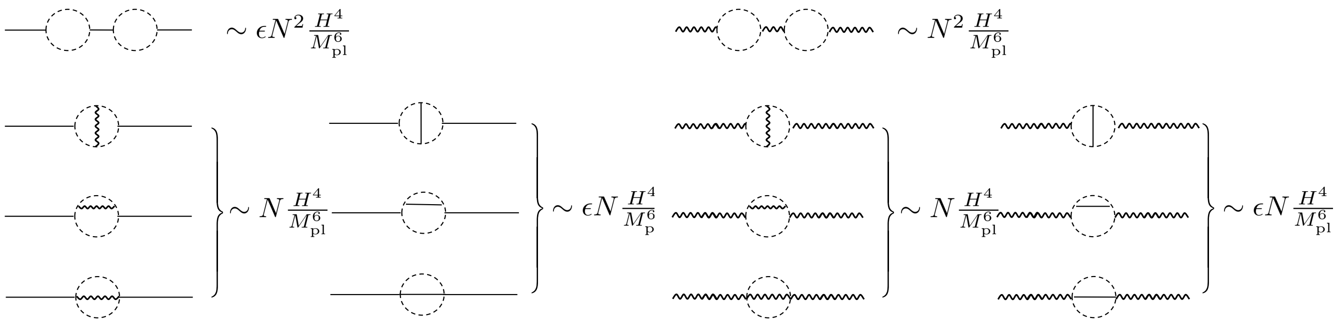

FIGURE 4.3. (a) Two loop contributions to the propagator $\langle\zeta\zeta\rangle$.Solid lines denote the curvature perturbation propagator; wavy lines represent the graviton propagator; while dash lines denote the external χ field. The double sunset graph dominates when $N >> 1/\epsilon$. (b) Two loop contributions to $\langle\gamma\gamma\rangle$, where now it is only necessary $N >> 1$ for the double sunset graph to dominate.

The same reasoning applies for the tensor spectrum and the corresponding resummation gives

$$P_{\gamma\, tot} = \frac{\frac{2H^2}{\pi^2 M_{pl}^2}\left(\frac{k}{k_*}\right)^{n_T}}{1 + N\epsilon \frac{H^2}{30 M_{pl}^2}\log\frac{k}{k_\star}}\,. \tag{4.41}$$

where $P_\gamma = \frac{2H^2}{\pi^2 M_{pl}^2}\left(\frac{k}{k_*}\right)^{n_T}$ is the predicted tree-level result for the tensor spectrum.

Logarithmic running

Referring to the parametrization used by the Planck collaboration [9]

$$\log P_\zeta = \log\Delta + \log k/k_* \left[n_s(k_*) - 1 + \frac{1}{2}\frac{d n_s}{d\log k}\bigg|_{k_*} \log k/k_* + \ldots \right] \tag{4.42}$$

we see from (Table 4) in the same reference that

$$n_s(k_\star) - 1 \;=\; \frac{d\log P_{tot}}{d(\log k)}\bigg|_{k=k_\star} \;=\; 0.9644 \pm 0.0049, \tag{4.43}$$

$$\alpha(k_\star) \;=\; \frac{d^2 \log P_{tot}}{d(\log k)^2}\bigg|_{k=k_\star} = -0.0085 \pm 0.0076, \tag{4.44}$$

$$r(k_\star) \;<\; 0.149\,. \tag{4.45}$$

where the pivot scale is taken to be $k_\star = 0.05\text{Mpc}^{-1}$. From Eqs. (4.40) and (4.41), one infers additional corrections to the tilt and the running:

$$n_s^{tot}(k) - 1 = n_s(k) - 1 - \frac{\frac{\epsilon^2 N H^2}{15 M_{pl}^2}}{1 + \frac{\epsilon^2 N H^2}{15 M_{pl}^2}\log(k/k_\star)} \tag{4.46}$$

$$\alpha_s^{tot}(k) = \alpha_s(k) + \alpha_N^s(k) = \alpha_s(k) + \frac{1}{\left[1 + N\frac{\epsilon^2 H^2}{15 M_{pl}^2}\log\frac{k}{k_\star}\right]^2}\left[\frac{\epsilon^2 N H^2}{15 M_{pl}^2}\right]^2 \tag{4.47}$$

$$r_{tot}(k) = r\left[\frac{1 + \epsilon^2 N \frac{4\pi}{15}\frac{H^2}{M_{pl}^2}\log(k/k_\star)}{1 + \frac{N\epsilon}{30}\frac{H^2}{M_{pl}^2}\log(k/k_\star)}\right] \tag{4.48}$$

$$n_T^{tot}(k) = n_T(k) - \frac{\frac{\epsilon N H^2}{30 M_{pl}^2}}{1 + \frac{\epsilon N H^2}{30 M_{pl}^2}\log(k/k_\star)} \tag{4.49}$$

$$\alpha_T^{tot}(k) = \alpha_T(k) + \alpha_N^T(k) = \alpha_T(k) + \frac{1}{\left[1 + N\frac{\epsilon H^2}{30 M_{pl}^2}\log\frac{k}{k_\star}\right]^2}\left[\frac{\epsilon N H^2}{30 M_{pl}^2}\right]^2 \tag{4.50}$$

To second order in the Hubble hierarchy one gets [56, 175]:

$$n_s^{tot}(k_\star) - 1 = -2\epsilon - \epsilon_2 - 2\epsilon^2 - (2C+3)\epsilon\epsilon_2 - C\epsilon_2\epsilon_3 + 2\epsilon^2\lambda_N\,, \tag{4.51}$$

$$\alpha_{tot}(k_\star) = -2\epsilon\epsilon_2 - \epsilon_2\epsilon_3\,, \tag{4.52}$$

$$r_{tot}(k_\star) = r(k_\star) = 16\epsilon \tag{4.53}$$

$$n_T^{tot}(k_\star) = -2\epsilon + \epsilon\lambda_N + 2\epsilon^2 + 2(C+1)\epsilon\epsilon_2\,, \tag{4.54}$$

$$\alpha_{T\,tot}(k_\star) = -\lambda_N^2 - 2\epsilon\epsilon_2\,. \tag{4.55}$$

Here $C = \gamma_E + \log 2 - 2$, $\gamma_E \simeq 0.577$ is the Euler-Mascheroni constant, and we have introduced

$$\lambda_N = N\frac{H^2}{30 M_{pl}^2} = N\frac{\pi^2 r_\star}{60}\Delta_\zeta < N \cdot 10^{-10}\,. \tag{4.56}$$

In this formulas we include the standard slow roll results up to second order in the slow roll parameters [56, 175].

From these results it is clear that the tensor spectral index and its running contain the clearest signature of the scalar field contributions. Because N does not appear in the leading order term in slow-roll parameters in the scalar tilt and running, its contribution is too feble to compete.

The leading order contribution to the spectral tilt of tensor modes becomes $n_T^{tot} = -2\epsilon + \lambda\epsilon$. The consistency relation is then modified since $r_\star$, as found above, is unchanged, but now

$$n_T^{tot} = -\frac{r_\star}{8}\left[1 - \frac{\lambda}{2}\right] \tag{4.57}$$

The value predicted for N using (4.56) is

$$N = \Delta_\zeta^{-1}\frac{15}{\pi^2 r_\star^2}\left[r_\star + 8n_T\right] \tag{4.58}$$

As a consequence, if we can bound the consistency relation by some positive number, i.e. if, to some degree of confidence we can find some upper bound ξ:

$$r_\star + 8n_T < \xi \tag{4.59}$$

then

$$N \lesssim 10^9 \frac{\xi}{r_\star^2} \tag{4.60}$$

The most optimistic case is when $r_\star \sim 10^{-1}$, which would imply: $N \lesssim 10^{11}\xi$. So the more the consistency relation is violated, the more stringent bound this calculation provides.

Comparison with strong coupling bounds

The above results are to be compared with the strong coupling bound, which states that gravity becomes strongly coupled at the scale M_{pl}^2/N [79]. This implies an upper bound on the curvature of spacetime before quantum gravity effects become relevant. Hence a consistent semi-classical treatment requires that $H^2 < M_{pl}^2/N$, implying that the number of species cannot be greater than

$$N < \frac{2\Delta_\zeta^{-1}}{\pi^2 r_\star} =: B\,, \qquad \lambda_N < 1\,. \tag{4.61}$$

We can rewrite now (4.58) as

$$N = \Delta_\zeta^{-1}\frac{15}{\pi^2 r_\star}\left[1 + \frac{8n_T^{tot}}{r_\star}\right] = B\frac{15}{2}\left[1 - \frac{-8n_T^{tot}}{r_\star}\right] \tag{4.62}$$

Therefore, deviations on the tensor to scalar consistency relation $\frac{-8n_T^{tot}}{r_\star} = 1$ improve the bound on N imposed by the strong coupling. The more deviated, the strongest constraint on N. Note also that, in order to improve the bound on N, it is crucial that the eventual measure of n_T^{tot} gives a negative value. Assuming this is the case, the experimental bound on $r_\star$ (4.45) leads to

$$N < B\frac{15}{2}[1-50|n_T|] \tag{4.63}$$

The greater the eventual measure of $|n_T|$ is, the stronger the bound on N will be. Conversely, given a particular value of N, predicted by a specific theory, one could bound the expected $|n_T|$ to be measured.

Of course the strong coupling bound is derived from theoretical consistency of the theory, while the bounds we have derived are experimental (assuming a phase of slow roll inflation has taken place).

4.4 Discussion

From the fact that even hidden quantum fields generate, at 1-loop, perturbations in the metric with a non-vanishing spectrum, we have derived a limit on the number of such scalar fields. For this we have demanded that their presence does not spoil inflation (the strong coupling bound). The discussion in this paper is easily extended to fields with higher spin or to non-minimally coupled scalar fields. All fields lead to very similar corrections the ζ and γ which differ mainly in their pre-factor but not in the parametric form. We therefore conclude that N can be considered as the total number of fundamental degrees of freedom with masses $m \ll H$ during inflation. This will be worked out in future investigations.

The bounds found in this work may put serious challenges to some phenomenological models that request the existence of a large amounts of light degrees of freedom in the universe. Some of these models propose to solve the radiative stability of hierarchy between the Planck and the weak (TeV) scale (so-called "Hierarchy problem") by assuming large numbers of particle species. For instance, in [80], it was proposed that 10^{32} mirror copies of the standard model coupled only through gravity would explain the problem. It is conjectured that the fundamental cutoff of gravity $M_\star$ could actually be TeV and that M_{pl} is a derived scale $M_{pl} \sim \sqrt{N} M_\star$. A similar situation was considered in [29], in which N copies of the standard model are introduced, each one with different values of the higgs mass. Depending on whether gauge coupling unification and SUSY is considered or not, they require the existence of 10^4 or up to 10^{16} extra degrees of freedom. The bounds considered in this work put additional arguments to constrain the parameter space.

4.5 Appendices

A. On the ϵ dependence of the vertices

We consider the action for the zero mode of the (canonically normalized) inflaton plus N hidden scalars.

$$\begin{aligned} S &= \frac{M_{pl}^2}{2}\int d^4x\sqrt{-g}R[g] - \frac{1}{2}\int d^4x\sqrt{-g}\left[\partial_\mu\phi\partial^\mu\phi + 2V(\phi)\right] \\ &- \sum_{n=1}^{N}\frac{1}{2}\int d^4x\sqrt{-g}\left[\partial_\mu\chi_n\partial^\mu\chi_n + m_n^2\chi_n^2\right] + \ldots \end{aligned} \tag{4.64}$$

Where the ellipses denote higher order terms which we will consider further. By assumption, the χ fields have no classically evolving background, and so appear in the action to leading order as quadratic in perturbations. We ADM decompose the metric

$$ds^2 = -N^2dt^2 + h_{ij}(dx^i + N^i dt)(dx^j + N^j dt), \tag{4.65}$$

and work in comoving gauge

$$\phi(t,x) = \phi_0(t), \tag{4.66}$$

$$h_{ij}(t,x) = a^2(t)e^{2\zeta(t,x)}\delta_{ij}. \tag{4.67}$$

This gauge is defined by the foliation where the inflaton is the clock (no other field has a background). Writing

$$\begin{aligned} N &= 1 + \alpha_1 \\ N^i &= \partial_i\theta + N_T^i\ ,\ w/\partial_i N_T^i \equiv 0 \end{aligned} \tag{4.68}$$

where α_1, θ and N_T^i all first order quantities, we find the solutions (we only need to calculate to first order for the constraints to obtain the action to cubic order [131])

$$\alpha_1 = \frac{\dot\zeta}{H} \tag{4.69}$$

$$\partial^2\theta = -\frac{\partial^2\zeta}{a^2H} + \epsilon\dot\zeta \tag{4.70}$$

where $\partial^2 \equiv \partial_i\partial_i$ contains no factors of the scale factor, and where ϵ is defined as:

$$\epsilon := \frac{\dot\phi_0^2}{2H^2M_{pl}^2} \tag{4.71}$$

The relevant quadratic and cubic terms are (summation over n implicit)

$$S_{2,\zeta} = M_{pl}^2 \int d^4x\, a^3 \epsilon \left[\dot{\zeta}^2 - \frac{1}{a^2}(\partial\zeta)^2\right] \tag{4.72}$$

$$S_{2,\chi} = \frac{1}{2}\int d^4x\, a^3 \left[\dot{\chi}_n\dot{\chi}_n - \frac{1}{a^2}\partial_i\chi_n\partial_i\chi_n - m_n^2\chi_n^2\right] \tag{4.73}$$

$$\begin{aligned} S_{3,\zeta\chi} = \frac{1}{2}\int d^4x \Big\{ &\quad a^3\dot{\chi}_n\dot{\chi}_n\left(3\zeta - \frac{\dot{\zeta}}{H}\right) - 2a^3\dot{\chi}_n\partial_i\theta\partial_i\chi_n \\ &- \quad a^3\left(\zeta + \frac{\dot{\zeta}}{H}\right)\frac{1}{a^2}\partial_i\chi_n\partial_i\chi_n - a^3\left(3\zeta + \frac{\dot{\zeta}}{H}\right)m_n^2\chi_n^2\Big\} \end{aligned} \tag{4.74}$$

We do not write the cubic action for ζ since all we will need from it is the fact that it is suppressed by ϵ^2 to leading order after enough integrations by parts [131]. A similar thing happens for (4.74) – although it may appear that the cubic interactions between ζ and the χ^a might be of order ϵ^0, these interactions in fact of order ϵ. This is readily seen by realizing that this contribution to the action is nothing other than the variation of the quadratic action for the hidden fields to first order in metric perturbations. That is, if $\mathscr{L}_\chi = -\frac{1}{2}(\partial_\mu\chi_n\partial^\mu\chi_n + m_n^2\chi_n^2)$, then the cubic interaction action for the hidden fields is given merely by the first order variation

$$S_{3,\zeta\chi} = \delta_{g_{\mu\nu}}\int d^4x\sqrt{-g}\,\mathscr{L}_\chi = \frac{1}{2}\int d^4x\sqrt{-g}\,T_\chi^{\mu\nu}\delta_1 g_{\mu\nu}, \tag{4.75}$$

where $T_\chi^{\mu\nu}$ corresponds to the stress-energy tensor of a minimally coupled massless scalar field in an unperturbed background

$$T_{\mu\nu}^\chi = -\frac{\bar{g}^{\mu\nu}}{2}\left(\partial_\lambda\chi_n\partial^\lambda\chi_n + m_n^2\chi_n^2\right) + \partial_\mu\chi_n\partial_\nu\chi_n \tag{4.76}$$

From (4.65), (4.69) and (4.70) we see that the first order metric variations can be read off as

$$\delta_1 g_{\mu\nu} = \begin{pmatrix} -2\frac{\dot{\zeta}}{H} & a^2\partial_i\theta \\ a^2\partial_i\theta & a^2\delta_{ij}2\zeta \end{pmatrix} \tag{4.77}$$

One can explicitly verify that the trace of the product of the above with (4.76) reproduces (4.74). We observe that one can write (4.77) as [6]

$$\delta_1 g_{\mu\nu} = \nabla_\mu\beta_\nu + \nabla_\nu\beta_\mu + \Delta_{\mu\nu}, \tag{4.78}$$

[6] the underlying reason of this decomposition lies in the fact that the unambigous physical contribution to scalar perturbations is obtained by writing down the choice of gauge explicitly as a time diffeomorphism, $x^\mu \to x^\mu + \beta^\mu$, where β^0 is the goldstone boson eaten by the metric in comoving gauge [57]

where

$$\beta_0 = -\frac{\zeta}{H}, \qquad \beta_i \equiv 0 \tag{4.79}$$

and where

$$\Delta_{\mu\nu} := \epsilon \begin{pmatrix} 2\zeta & a^2\partial_i\partial^{-2}\dot{\zeta} \\ a^2\partial_i\partial^{-2}\dot{\zeta} & 0 \end{pmatrix} \tag{4.80}$$

Because $\nabla_\mu T^{\mu\nu}_\chi = 0$, clearly only the second term in (4.78) gives a non-vanishing contribution. Therefore the relevant cubic interactions are given by

$$S_{3,\zeta\chi} = \int d^4x a^3 \epsilon \left[\frac{\zeta}{2}\left(\dot{\chi}_n \dot{\chi}^n + \frac{1}{a^2}\partial_i\chi_n\partial_i\chi_n + m_n^2\chi_n^2 \right) - \dot{\chi}_n\partial_i\chi_n\partial_i\partial^{-2}\dot{\zeta} \right] \tag{4.81}$$

where we take note of the advertised ϵ suppression of the cubic interaction vertices.

Similarly, the relevant cubic interaction term for the tensor perturbations can also be written as (4.75), but now

$$\delta_1 g_{\mu\nu} = \begin{pmatrix} 0 & 0 \\ 0 & a^2\gamma_{ij} \end{pmatrix} \tag{4.82}$$

which has no gauge ambiguities. Thus

$$S_{3,\gamma\chi} = \frac{1}{2}\int d^4x a^5 T^{ij}\gamma_{ij} = \frac{1}{2}\int d^4x a^5 \Pi^{ij}\gamma_{ij} \tag{4.83}$$

where in the last line we used the fact that γ_{ij} selects the traceless and transverse contribution of the stress-energy tensor.

B. Equation of motion for curvature perturbations during inflation

We shall approach the problem now from the perspective of the equations of motion We derive a second-order differential equation determining the evolution of the comoving curvature perturbation ζ in presence of external sources such as the N hidden scalar fields, following closely to [50]. This will result in a generalized Mukhanov-Sasaki equation. As we shall see, energy momentum conservation will be very crucial for deriving the correct result, as it was in getting from (4.74) to (4.81) with the aim of determining the correct order of ϵ of the vertex interaction. [7]

[7] The main difference with [50] is that in our case the scalar field energy momentum tensor is separately conserved while in [50] a $U(1)$ gauge field is considered which is non-minimally coupled to the inflaton. The resulting equations of motion are different. This is why a new derivation needs to be done.

The background inflaton field ϕ_0 with its slow-roll potential $V(\phi_0)$, satisfies the equation of motion $\ddot{\phi}_0 + 3H\dot{\phi}_0 + V'(\phi_0) = 0$. Its energy-momentum tensor reads

$$T_0^{\mu\nu} \;=\; -\bar{g}^{\mu\nu}\left[\frac{1}{2}\partial_\sigma\phi_0\partial^\sigma\phi_0 + V(\phi_0)\right] + \partial^\mu\phi_0\partial^\nu\phi_0 \,. \qquad (4.84)$$

It has the form of the energy-momentum tensor of a perfect fluid, $T^{\mu\nu} = (\rho+p)u^\mu u^\nu + p\bar{g}^{\mu\nu}$ with $u^\mu = \frac{1}{\sqrt{-\partial_\sigma\phi_0\partial^\sigma\phi_0}}\partial^\mu\phi_0$ for the energy-flux vector. Since $u^\mu u_\mu = -1$, one can choose a foliation of spacetime with this unit timelike vector, and write $\bar{g}_{\mu\nu} = -u_\mu u_\nu + h_{\mu\nu}$. Given that the field is homogeneous, $\phi_0 = \phi_0(t)$, this vector is simply $u^\mu = (-1,0,0,0)$ in the FLRW coordinate basis. In other words, ϕ_0 is our "clock". Energy density and pressure are given by

$$\rho_0 \;=\; u_\mu u_\nu T_0^{\mu\nu} = -\frac{1}{2}\partial_\sigma\phi_0\partial^\sigma\phi_0 + V(\phi_0)\,, \qquad (4.85)$$

$$p_0 \;=\; \frac{1}{3}h_{\mu\nu}T_0^{\mu\nu} = -\frac{1}{2}\partial_\sigma\phi_0\partial^\sigma\phi_0 - V(\phi_0)\,. \qquad (4.86)$$

We now introduce an external source consisting of a set of N minimally-coupled light scalar fields $\{\chi_n\}$. The energy-momentum tensor is again given by (4.84) with a negligible potential, but in the foliation determined by u_μ, and since we admit the χ-fields to be inhomogeneous, this no longer has the form of a perfect fluid, and non-vanishing anisotropic stresses $\Pi^{\mu\nu}$ and energy fluxes q^μ will appear. The corresponding decomposition in the foliation defined by ϕ_0 reads

$$T_\chi^{\mu\nu} = (\rho+p)u^\mu u^\nu + p\bar{g}^{\mu\nu} + 2u^{(\mu}q^{\nu)} + \Pi^{\mu\nu}\,. \qquad (4.87)$$

The imposition, $\nabla_\mu T_\chi^{\mu\nu} = 0$, leads to 2 conservation equations of the scalar sector:

$$0 \;=\; \dot{\rho} + 3H(\rho+p) - \partial_i q^i\,, \qquad (4.88)$$

$$0 \;=\; \partial_i\dot{q}^i + 3H\partial_i q^i - \frac{\partial^2}{a^2}p + \frac{2}{3a^2}\partial^2\Pi_s\,, \qquad (4.89)$$

where $\partial^2 \equiv \partial_i\partial_i$, $q^i = -u_\mu T^{\mu i}$, and the scalar contribution of the anisotropic stress is parametrized by $\Pi^s_{ij} = -a^2(\partial^{-2}\partial_i\partial_j - \frac{1}{3}\delta_{ij})\Pi_s$. In Fourier space we introduce the energy flux potential q via

$$k^2 q = \partial_i q_i = ik_i q_i\,. \qquad (4.90)$$

Equivalently the scalar anisotropic stress potential is given by

$$\frac{2}{3}k^2\Pi_s = a^2(-k_ik_j + \frac{k^2}{3}\delta_{ij})T^{ij}\,. \qquad (4.91)$$

Note that q has dimension 3 while all other variables describing the χ-energy momentum tensor have dimension 4. Using conformal time, the conservation equations in Fourier space simply become

$$0 = \rho' + 3\mathcal{H}(\rho + p) - \frac{k^2}{a} q\,, \tag{4.92}$$

$$0 = q' + 3\mathcal{H} q + p - \frac{2}{3}\Pi_s\,. \tag{4.93}$$

We have to consider inhomogeneous linear perturbations of the inflaton field by setting $\phi(t,\mathbf{x}) = \phi_0(t) + \delta\phi(t,\mathbf{x})$, which are responsible for the primordial density fluctuations in the early universe. As a result, this leads to a perturbed metric $g_{\mu\nu}$. We shall work in the unitary gauge, defined by $\delta\phi(t,\mathbf{x}) = 0$. As an additional gauge condition we set $g_{0i} \equiv 0$. This (nearly) completely fixes the gauge. The first-order perturbed FLRW metric can then be written as

$$g_{\mu\nu} dx^\mu dx^\nu = a^2 \left\{ -(1+2A)d\eta^2 + \left[\left(1 + 2H_L + 2\frac{H_T}{3} \right) \delta_{ij} + 2H_T Y_{,ij} \right] dx^i dx^j \right\}. \tag{4.94}$$

The first order perturbations of the inflaton energy-momentum tensor in this gauge are (in Fourier space)

$$\delta\rho_\phi = \dot\phi_0 \delta\dot\phi + V'(\phi)\delta\phi - A\dot\phi_0^2 = -A\dot\phi_0^2\,, \tag{4.95}$$

$$\delta p_\phi = \dot\phi_0 \delta\dot\phi - V'(\phi)\delta\phi - A\dot\phi_0^2 = -A\dot\phi_0^2\,, \tag{4.96}$$

$$\delta q_\phi := i\frac{k_i}{k^2}\delta q_i^\phi = i\frac{k_i}{k^2} i k_i \dot\phi_0 \delta\phi = -\dot\phi_0 \delta\phi = 0\,, \tag{4.97}$$

Inserting this in the first-order perturbed Einstein equations $\delta G_{\mu\nu} = M_{pl}^{-2}\delta T_{\mu\nu}$ we find

$$3\mathcal{H}^2 A - 3\mathcal{H} H_L' - k^2 (H_L + \frac{H_T}{3}) = \frac{a^2}{2M_{Pl}^2}\left\{ \frac{\phi_0'^2}{a^2} A - \rho \right\} \tag{4.98}$$

$$\mathcal{H} A - H_L' - \frac{H_T'}{3} = -\frac{a}{2M_{pl}^2} q\,, \tag{4.99}$$

$$H_T'' + 2\mathcal{H} H_T' - k^2 (A + H_L + \frac{H_T}{3}) = \frac{a^2}{M_{Pl}^2}\Pi_s\,, \tag{4.100}$$

$$\mathcal{H} A' + (2\mathcal{H}' + \mathcal{H}^2) A - \frac{k^2}{3}(A + H_L + \frac{H_T}{3}) - 2\mathcal{H} H_L' - H_L'' = \frac{a^2}{2M_{pl}^2}\left\{ -\frac{\phi_0'^2}{a^2} A + p \right\} \tag{4.101}$$

The first two are the constraint equations, while (4.100) and (4.100) are the evolution equations. Here ρ, p, Π_s, and $q_i \equiv -k_i q$ represent the energy density, pressure, scalar anisotropic stress, and energy-flux vector of the external source, $\{\chi_n\}$. Primes denote derivative with respect conformal time, $d\eta = \frac{dt}{a}$. The equations above have been extracted from [50], and they have been checked with Appendix B of [43] by translating the notation.

The comoving curvature perturbation ζ is gauge invariant, and hence a suitable physical quantity to describe scalar cosmological perturbations. In our gauge it is given by $\zeta = +H_L + \frac{H_T}{3} - \frac{H}{\rho_0+p_0}\delta q_\varphi = +H_L + \frac{H_T}{3}$ The above set of equations can be rearranged so to obtain an evolution equation for ζ:

$$(\mathcal{H} - \frac{\mathcal{H}'}{\mathcal{H}})\zeta'' + \left[2\mathcal{H}^2 - 2\mathcal{H}' + \left(\frac{\mathcal{H}'}{\mathcal{H}}\right)^2 - \left(\frac{\mathcal{H}'}{\mathcal{H}}\right)'\right]\zeta' + (\mathcal{H} - \frac{\mathcal{H}'}{\mathcal{H}})k^2\zeta = \tag{4.102}$$
$$+\frac{a^2}{2M_{pl}^2}\left\{\rho' - \frac{q'}{a}\left[2\mathcal{H} + \frac{\mathcal{H}'}{\mathcal{H}}\right] - \rho\left[\frac{\mathcal{H}'}{\mathcal{H}} - 4\mathcal{H}\right]\right.$$
$$\left. - \frac{q}{a}\left[k^2 + \mathcal{H}' - 2\mathcal{H}^2 + \left(\frac{\mathcal{H}'}{\mathcal{H}}\right)' - \left[2\mathcal{H} + \frac{\mathcal{H}'}{\mathcal{H}}\right]\left[\frac{\mathcal{H}'}{\mathcal{H}} - 4\mathcal{H}\right]\right] + 2\mathcal{H}\Pi_s\right\}.$$

We make use of the following identities which are realized in slow-roll inflation:

$$\mathcal{H}^2 - \mathcal{H}' = \epsilon\mathcal{H}^2, \qquad \frac{\phi_0''}{\phi_0'} = \mathcal{H}(1+\delta), \tag{4.103}$$

where ϵ and δ are the slow roll parameters, defined in (4.6). We also use [189] $\dot{\epsilon} = 2H\epsilon[\delta+\epsilon]$. At leading leading order in slow-roll parameters, we find

$$\zeta'' + \left[2\mathcal{H} + \frac{\epsilon'}{\epsilon}\right]\zeta' + k^2\zeta \quad = \quad \frac{+a^2}{2M_{pl}^2\epsilon\mathcal{H}}\left\{\rho' - \frac{q'}{a}\mathcal{H}(3-\epsilon) + \rho\mathcal{H}(3+\epsilon)\right.$$
$$\left. - \frac{q}{a}\left(k^2 + 9\mathcal{H}^2 - 3\epsilon\mathcal{H}^2 - \epsilon'\mathcal{H}\right) + 2\mathcal{H}\Pi_s\right\}.$$

The leading order contribution on the RHS is actually of order ϵ^0, and not ϵ^{-1}. Using the conservation equations (4.92) and (4.93), one finds that all the terms of order ϵ^{-1} cancel and we remain with

$$\zeta'' + 2\frac{z'}{z}\zeta' + k^2\zeta = \frac{a^2}{2M_{pl}^2}\left\{\rho - p + \frac{2}{3}\Pi_s + \frac{q}{a}\frac{\epsilon'}{\epsilon}\right\}. \tag{4.104}$$

which is the Mukhanov-Sasaki equation, with a source term determined by the scalar fields χ_n. It is customary to write it in terms of $z^2 \equiv a^2 2M_{pl}^2\epsilon = a^2\frac{\phi_0^2}{H^2}$. Note that it is important that the hidden fields are separately conserved (do not interact with the inflaton field) in order for the ϵ^{-1} contributions to the source term to vanish. If this were not the case, we would obtain much larger corrections and therefore much tighter constraints.

The correction for the power spectrum will have an order of magnitude ϵ in agreement with the findings of [188]. The equation is manifestly gauge-invariant, and is not limited to scalar fields, since in the derivation we made only use of the general form (4.87) of the energy-momentum tensor. It can also be applied to other matter fields, Dirac spinors for instance. Note also that in the non-relativistic limit the Poisson equation $+\frac{k^2}{a^2}\zeta = 4\pi G\rho$ is recovered. Expression (4.104) is our central equation. The last term can actually be ignored as it is of higher order in the slow-roll parameters.

We now show that (4.104) can also be obtained via the action. The relevant cubic interaction in the action derived in (4.81) can be arranged as

$$\begin{aligned} S_{3,\zeta\chi} &= \int d^4x\, a^3\epsilon\left\{\frac{\zeta}{2}\left[\dot\chi\dot\chi + \frac{1}{a^2}\partial_i\chi\partial_i\chi\right] + (\partial^{-2}\dot\zeta)\partial_i(\dot\chi\partial_i\chi)\right\} && (4.105)\\ &= \int d^4x\, a^3\epsilon\left\{\frac{\zeta}{2}\left[\dot\chi\dot\chi + \frac{1}{a^2}\partial_i\chi\partial_i\chi\right] + \dot\zeta\partial^{-2}\partial_i(\dot\chi\partial_i\chi)\right\} && (4.106)\\ &= \int d^4x\, a^3\epsilon\{\zeta\rho - q\dot\zeta\} = \int d^4x\, a^3\epsilon\,\zeta\left[\rho + \dot q + 3Hq + \frac{\dot\epsilon}{\epsilon}q\right] && (4.107)\end{aligned}$$

Using momentum conservation to eliminate $\dot q + 3Hq$ we obtain the source term $\rho + \dot q + 3Hq + \frac{\dot\epsilon}{\epsilon}q = \rho - p + \frac{2}{3}\Pi_s + \frac{\epsilon'}{\epsilon}\frac{q}{a}$, which is in full agreement with (4.104).

C. Contribution from the hidden sector to the power spectra

In this appendix we want to determine the particular solution of (4.104) due to the source, correcting the wellknown homogeneous result. We present the details of the calculation of the two-point function.

For convenience we introduce an auxiliary field $\varphi_k(\eta) = z\zeta_k(\eta)$. In terms of φ_k equation (4.104) simplifies to

$$\left[\frac{d^2}{d\eta^2} + k^2 - \frac{1}{z}\frac{d^2z}{d\eta^2}\right]\varphi_k = \sqrt{\frac{\epsilon}{4\pi G}}\frac{4\pi G}{H^3}\frac{S_k(\eta)}{\eta^3}. \tag{4.108}$$

where $S_k(\eta)$ can be read off from (4.104). The general inhomogeneous solution of this equation is given by (we assume the slow-roll field approximation, in which ϵ, δ are constant with time)

$$\varphi_k(\eta) = \sqrt{\frac{\epsilon}{4\pi G}}\frac{4\pi G}{H^3}\int_{-\infty}^{0} d\eta' \mathscr{G}_k^R(\eta,\eta')\frac{S_k(\eta')}{\eta'^3}, \tag{4.109}$$

where $\mathscr{G}_k^R(\eta,\eta')$ is the retarded Green function of (4.108). Since $\mathscr{G}_k^R(\eta,\eta') \propto \Theta(\eta-\eta') = 0$ for $\eta < \eta'$, the homogeneous (Mukhanov-Sasaki) solution corresponds to the solution of the field equation at $\eta \to -\infty$, i.e. before the interaction with the external source has taken place.

Vacuum expectation values of products of free fields can be associated with the Green functions of the equation of motion [48]. The so-called Hadamard elementary function (Feynman propagator) $\mathcal{G}(x,x')$ of a linear differential equation like the Klein-Gordon equation is the difference of the advanced and retarded Green functions, $\mathcal{G}_R(x,x') = -\Theta(\eta-\eta')\mathcal{G}(x,x')$ and $\mathcal{G}_A(x,x') = \Theta(\eta'-\eta)\mathcal{G}(x,x')$. It appears as the vacuum expectation value of the anticommutator of a scalar field with itself,

$$\mathcal{G}(x,x') = \langle\{\varphi(x),\varphi(x')\}\rangle = \int \frac{d^3k}{(2\pi)^3}\mathcal{G}_k(\eta,\eta')e^{i\mathbf{k}(\mathbf{x}-\mathbf{x}')}. \tag{4.110}$$

which for (exact) de Sitter space leads to, $y_k \equiv -k\eta$, $y'_k \equiv -k\eta'$

$$\mathcal{G}_k(\eta,\eta') = 2\mathcal{R}e\left[\varphi_k(\eta)\varphi^*_k(\eta')\right] = \frac{1}{k y_k y'_k}\left[(1+y_k y'_k)\cos(y_k-y'_k)+(y_k-y'_k)\sin(y_k-y'_k)\right] \tag{4.111}$$

Up to a normalization factor y'^3 which must be included, the corresponding retarded function is indeed a Green function:

$$\begin{aligned}\left[\frac{d^2}{d\eta^2}+k^2-\frac{1}{z}\frac{d^2z}{d\eta^2}\right](-1)\mathcal{G}_k(\eta,\eta')\Theta(\eta-\eta') &= -k\left[\frac{d^2}{dy_k^2}+1-\frac{2}{y_k^2}\right]\mathcal{G}_k(y_k,y'_k)\Theta(\eta-\eta') = \frac{k}{y'^3_k}\delta(y_k-y'_k)\\ &= \frac{1}{y'^3_k}\delta(\eta-\eta')\end{aligned}$$

and hence, the formal expression (4.109) is a solution of the inhomogeneous equation (4.108) if we set $\mathcal{G}^R_k(\eta,\eta') = -y'^3_k\mathcal{G}_k(\eta,\eta')\Theta(\eta-\eta')$. For practical purposes in future calculations, it is convenient to take the superhorizon limit of (4.111): $y_k \ll 1$ or $y'_k \ll 1$. It yields,

$$y'^3_k\mathcal{G}_k(\eta,\eta') \approx \frac{\eta'^2}{\eta}, \quad y'_k \ll 1, \text{and } y_k \ll 1. \tag{4.112}$$

For a slow-roll inflationary background, the modes become $\varphi_k(\eta) = \frac{\sqrt{-\pi\eta}}{2}H^{(1)}_{3/2+2\epsilon+\delta}(-k\eta)$, the large-scale behaviour of the retarded Green function is given by

$$\mathcal{G}^R_k(\eta,\eta') := (-y'^3_k)\mathcal{G}_k(\eta,\eta')\Theta(\eta-\eta') \approx -\frac{\eta'^{2+2\epsilon+\delta}}{\eta^{1+2\epsilon+\delta}}\Theta(\eta-\eta'), \quad y'_k << 1, \text{and } y_k << 1 \tag{4.113}$$

We use this expression to determine the inhomogeneous solution (4.109). But first, we need to determine the source which comes entirely from the hidden sector, the N external fields.

The equation of motion for a massless minimally coupled scalar field χ is given by $\Box\chi(x) = 0$. In order to solve the equation of motion, we descompose the field in terms of Fourier modes $\chi_{\mathbf{k}}$. We take advantage of the maximal group of spatial symmetries to decouple the temporal dependence from the spatial one, $\chi_{\mathbf{k}} = \chi_k(t)e^{i\mathbf{k}\cdot\mathbf{x}}$. The Klein-Gordon equation is translated then to a second-order differential equation of motion for the modes

$$\ddot{\chi}_k + 3H\dot{\chi}_k + \frac{k^2}{a^2}\chi_k = 0, \quad \chi_k'' + 2\mathscr{H}\chi_k' + k^2\chi_k = 0 \tag{4.114}$$

Assuming the natural extension of the Bunch-Davies vacuum state to slow-roll inflation, the standard mode solutions are

$$\chi_k(t) = \frac{\sqrt{-\pi\eta}}{2a}H_\mu^{(1)}\left(-k\eta\right). \tag{4.115}$$

As above, we use the variable $y_k \equiv -k\eta = \frac{k}{aH(1-\epsilon)}$ below.

With the modes solution (4.115) we now determine the source term in the inhomogeneous solution (4.109). In the foliation defined by the inflaton field ϕ_0, the energy-momentum tensor of an external field is of the form (4.87). A scalar field produces $T^\chi_{\mu\nu} = -\frac{1}{2}\bar{g}_{\mu\nu}\partial_\sigma\chi\partial^\sigma\chi + \partial_\mu\chi\partial_\nu\chi$. According to (4.104) we are interested in the following components:

$$\rho = T^\chi_{\mu\nu}u^\mu u^\nu = \frac{1}{2}\left[\dot{\chi}^2 + \frac{1}{a^2}(\nabla\chi)^2\right], \tag{4.116}$$

$$p = \frac{1}{3}h_{\mu\nu}T^{\mu\nu}_\chi = \frac{1}{2}\left[\dot{\chi}^2 - \frac{1}{3a^2}(\nabla\chi)^2\right], \tag{4.117}$$

$$\Pi_{ij} = h^\alpha_i h^\beta_j T^\chi_{\alpha\beta} - \frac{1}{3}T^\chi_{\alpha\beta}h^{\alpha\beta}h_{ij} = (\partial_i\chi)(\partial_j\chi) - \frac{1}{3a^2}(\nabla\chi)^2 h_{ij}. \tag{4.118}$$

The corresponding modes are obtained by applying the convolution theorem,

$$\begin{aligned}\rho_k &= \frac{1}{2}\left[(\dot{\chi}*\dot{\chi})_{\mathbf{k}} + \frac{1}{a^2}(\nabla\chi*\nabla\chi)_{\mathbf{k}}\right] = \frac{1}{2}\int\frac{d^3p}{(2\pi)^3}\left[(\dot{\chi})_{\mathbf{p}}(\dot{\chi})_{\mathbf{k-p}} + \frac{1}{a^2}(\vec{\nabla}\chi)_{\mathbf{p}}(\vec{\nabla}\chi)_{\mathbf{k-p}}\right] \\ &= \frac{1}{2}\int\frac{d^3p}{(2\pi)^3}\left[\dot{\hat{\chi}}_{\mathbf{p}}\dot{\hat{\chi}}_{\mathbf{k-p}} - \frac{\mathbf{p}(\mathbf{k}-\mathbf{p})}{a^2}\hat{\chi}_{\mathbf{p}}\hat{\chi}_{\mathbf{k-p}}\right].\end{aligned} \tag{4.119}$$

$$\begin{aligned}p_k &= \frac{1}{2}\left[(\dot{\chi}*\dot{\chi})_{\mathbf{k}} - \frac{1}{3a^2}(\vec{\nabla}\chi*\vec{\nabla}\chi)_{\mathbf{k}}\right] = \frac{1}{2}\int\frac{d^3\mathbf{p}}{(2\pi)^3}\left[(\dot{\chi})_{\mathbf{p}}(\dot{\chi})_{\mathbf{k-p}} - \frac{1}{3a^2}(\vec{\nabla}\chi)_{\mathbf{p}}(\vec{\nabla}\chi)_{\mathbf{k-p}}\right] \\ &= \frac{1}{2}\int\frac{d^3p}{(2\pi)^3}\left[\dot{\hat{\chi}}_{\mathbf{p}}\dot{\hat{\chi}}_{\mathbf{k-p}} + \frac{\mathbf{p}(\mathbf{k}-\mathbf{p})}{3a^2}\hat{\chi}_{\mathbf{p}}\hat{\chi}_{\mathbf{k-p}}\right].\end{aligned} \tag{4.120}$$

$$\begin{aligned}(\Pi_{ij})_{\mathbf{k}} &= (\partial_i\chi*\partial_j\chi)_{\mathbf{k}} - \frac{1}{3a^2}(\nabla\chi*\nabla\chi)_{\mathbf{k}}h_{ij} = \int\frac{d^3\mathbf{p}}{(2\pi)^3}\left[(\partial_i\chi)_{\mathbf{p}}(\partial_j\chi)_{\mathbf{k-p}} - \frac{1}{3a^2}(\vec{\nabla}\chi)_{\mathbf{p}}(\vec{\nabla}\chi)_{\mathbf{k-p}}h_{ij}\right] \\ &= -\int\frac{d^3p}{(2\pi)^3}\left[p_i(k_j-p_j) - \frac{1}{3a^2}\mathbf{p}(\mathbf{k}-\mathbf{p})h_{ij}\right]\hat{\chi}_{\mathbf{p}}\hat{\chi}_{\mathbf{k-p}}.\end{aligned} \tag{4.121}$$

where $\hat{\chi}_{\mathbf{k}}(t) = \hat{a}_{\mathbf{k}}\chi_k(t) + \hat{a}^\dagger_{-\mathbf{k}}\chi^*_k(t)$ is the Fourier transform of the external scalar field, $\chi(x) = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\hat{\chi}_{\mathbf{k}}(t)e^{i\mathbf{k}\cdot\mathbf{x}}$ (reality of $\chi(x)$ implies $\hat{\chi}^\dagger_{\mathbf{k}}(t) = \hat{\chi}_{-\mathbf{k}}(t)$). The scalar anisotropic contribution is defined by the relation $(\Pi_{ij})^s_{\mathbf{k}} \equiv -a^2\left(\hat{k}_i\hat{k}_j - \frac{1}{3}\delta_{ij}\right)\Pi^s_{\mathbf{k}}$; thus, $\Pi^s_{\mathbf{k}} = -\frac{3}{2}a^2\hat{k}^i\hat{k}^j(\Pi_{ij})_{\mathbf{k}}$. Namely,

$$\Pi^s_{\mathbf{k}} \quad = \quad \frac{3}{2}\int\frac{d^3\mathbf{p}}{(2\pi)^3}\left[\frac{\mathbf{k}\cdot\mathbf{p}}{a^2}\left(1-\frac{\mathbf{k}\cdot\mathbf{p}}{k^2}\right) - \frac{\mathbf{p}(\mathbf{k}-\mathbf{p})}{3a^2}\right]\hat{\chi}_{\mathbf{p}}\hat{\chi}_{\mathbf{k}-\mathbf{p}}.. \tag{4.122}$$

Taking into account all these expressions the source term in (4.104) reads

$$S_{\mathbf{k}} \equiv \rho_{\mathbf{k}} - p_{\mathbf{k}} + \frac{2}{3}\Pi^s_{\mathbf{k}} = \int\frac{d^3\mathbf{p}}{(2\pi)^3}\frac{p^2-(\hat{\mathbf{k}}\cdot\mathbf{p})^2}{a^2}\hat{\chi}_{\mathbf{p}}\hat{\chi}_{\mathbf{k}-\mathbf{p}} = \int\frac{d^3\mathbf{p}}{(2\pi)^3}\frac{p^2\sin^2\theta}{a^2}\hat{\chi}_{\mathbf{p}}\hat{\chi}_{\mathbf{k}-\mathbf{p}}. \tag{4.123}$$

D. Scalar power spectrum

The contribution of the source to the power spectrum of curvature perturbations is determined by the two-point function of modes (4.109),

$$k^3\langle\zeta_{\mathbf{k}}(\eta)\cdot\zeta_{\mathbf{k}'}(\eta)\rangle = (2\pi)^3P_\zeta(\eta)\delta^{(3)}(\mathbf{k}+\mathbf{k}')\,. \tag{4.124}$$

Thus,

$$\langle\zeta_{\mathbf{k}}(\eta)\cdot\zeta_{\mathbf{k}'}(\eta)\rangle \quad = \quad \langle\frac{\varphi_{\mathbf{k}}(\eta)}{z}\cdot\frac{\varphi_{\mathbf{k}'}(\eta)}{z}\rangle = \frac{1}{4M^4_{pl}a^2H^6}\int_{-\infty}^0 d\tau_1\int_{-\infty}^0 d\tau_2\frac{\mathcal{G}^R_k(\eta,\tau_1)}{\tau_1^3}\frac{\mathcal{G}^R_{k'}(\eta,\tau_2)}{\tau_2^3}\langle S_{\mathbf{k}}(\tau_1)\cdot S_{\mathbf{k}'}(\tau_2)\rangle\,.$$

We define the auxiliary power spectrum $\langle S_{\mathbf{k}}(\tau_1)\cdot S_{\mathbf{k}'}(\tau_2)\rangle = (2\pi)^3\mathcal{P}_S(\tau_1,\tau_2)\delta^{(3)}(\mathbf{k}+\mathbf{k}')$ which we first determine. The computation of $\langle S_{\mathbf{k}}(\tau_1)\cdot S_{\mathbf{k}'}(\tau_2)\rangle$ requires to study the vacuum expectation value of four fields, and, with the help of $\hat{\chi}_{\mathbf{k}}(t) = \hat{a}_{\mathbf{k}}\chi_k(t) + \hat{a}^\dagger_{-\mathbf{k}}\chi^*_k(t)$, only the following contributions do not vanish,

$$\begin{aligned}\langle\hat{\chi}_{\mathbf{p}}(\tau_1)\hat{\chi}_{\mathbf{k}-\mathbf{p}}(\tau_1)\hat{\chi}_{\mathbf{p}'}(\tau_2)\hat{\chi}_{\mathbf{k}'-\mathbf{p}'}(\tau_2)\rangle \quad &= \quad \langle a_{\mathbf{p}}a_{\mathbf{k}-\mathbf{p}}a^\dagger_{-\mathbf{p}'}a^\dagger_{-\mathbf{k}'+\mathbf{p}'}\rangle\,\chi_p(\tau_1)\chi_{k-p}(\tau_1)\chi^*_{p'}(\tau_2)\chi^*_{k'-p'}(\tau_2)\\ &+\langle a_{\mathbf{p}}a^\dagger_{-\mathbf{k}+\mathbf{p}}a_{\mathbf{p}'}a^\dagger_{-\mathbf{k}'+\mathbf{p}'}\rangle\,\chi_p(\tau_1)\chi^*_{k-p}(\tau_1)\chi_{p'}(\tau_2)\chi^*_{k'-p'}(\tau_2),\end{aligned} \tag{4.125}$$

using standard conmutation relations and $a_{\mathbf{p}}|0\rangle = 0$, we find

$$\begin{aligned}&\langle\hat{\chi}_{\mathbf{p}}(\tau_1)\hat{\chi}_{\mathbf{k}-\mathbf{p}}(\tau_1)\hat{\chi}_{\mathbf{p}'}(\tau_2)\hat{\chi}_{\mathbf{k}'-\mathbf{p}'}(\tau_2)\rangle =\\ &\qquad(2\pi)^6\left[\delta^{(3)}(\mathbf{k}-\mathbf{p}+\mathbf{p}')\delta^{(3)}(\mathbf{p}+\mathbf{k}'-\mathbf{p}') + \delta^{(3)}(\mathbf{p}+\mathbf{p}')\delta^{(3)}(\mathbf{k}-\mathbf{p}+\mathbf{k}'-\mathbf{p}')\right]\\ &\qquad\qquad\times\chi_p(\tau_1)\chi_{k-p}(\tau_1)\chi^*_{p'}(\tau_2)\chi^*_{k'-p'}(\tau_2)\\ &\quad+(2\pi)^6\delta(\mathbf{k})\delta(\mathbf{k}')\chi_p(\tau_1)\chi^*_{k-p}(\tau_1)\chi_{p'}(\tau_2)\chi^*_{k'-p'}(\tau_2).\end{aligned} \tag{4.126}$$

The first two terms are actually identical and integration over $d^3\mathbf{p}'$ yields

$$\begin{aligned}\mathscr{P}_S(\tau_1,\tau_2) &= 2\int \frac{d^3\mathbf{p}}{(2\pi)^3} M_{\mathbf{p},\mathbf{k}-\mathbf{p}}(\tau_1) M^*_{\mathbf{k}-\mathbf{p},\mathbf{p}}(\tau_2) \\ &+ (2\pi)^3\delta^{(3)}(\mathbf{k})\int \frac{d^3\mathbf{p}}{(2\pi)^3} N_p(\tau_1)\int \frac{d^3\mathbf{p}'}{(2\pi)^3} N_{p'}(\tau_2), \end{aligned} \tag{4.127}$$

where we have introduced the auxiliary functions (note the difference with (4.123): χ is a function while $\hat{\chi}$ is an operator)

$$M_{\mathbf{p},\mathbf{k}-\mathbf{p}} = \frac{p^2\ \sin^2\theta}{a^2}\chi_{\mathbf{p}}\chi_{\mathbf{k}-\mathbf{p}}, \tag{4.128}$$

$$N_p = \frac{p^2\ \sin^2\theta}{a^2}|\chi_{\mathbf{p}}|^2. \tag{4.129}$$

We neglect the contribution with $\delta^{(3)}(\mathbf{k})$ which is a zero mode and only contributes to the background. The power spectrum of curvature perturbations (4.124) yields finally

$$\begin{aligned} P_\zeta(\eta) &= \frac{\eta^2 k^3}{4M_{pl}^4 H^4}\int_{\eta_{in}}^0 d\tau_1 \int_{\eta_{in}}^0 d\tau_2 \frac{\mathscr{G}_k^R(\eta,\tau_1)}{\tau_1^3}\frac{\mathscr{G}_k^R(\eta,\tau_2)}{\tau_2^3}\mathscr{P}_S(\tau_1,\tau_2) \\ &= \frac{\eta^2 k^3}{2M_{pl}^4 H^4}\int \frac{d^3\mathbf{p}}{(2\pi)^3}\left|\int_{\eta_{in}}^0 d\eta' \frac{\mathscr{G}_k^R(\eta,\eta')}{\eta'^3} M_{\mathbf{p},\mathbf{k}-\mathbf{p}}(\eta')\right|^2. \end{aligned} \tag{4.130}$$

On subhorizon scales the mode functions oscillate rapidly and do not contribute significantly to the result, Therefore we have restricted the integrals to times which are between η_{in} and η, for which the sourcing modes have exited the horizon. This means,

$$Max\left[-\frac{1}{p},-\frac{1}{|\mathbf{k}-\mathbf{p}|}\right] < \eta' < \eta. \tag{4.131}$$

Equation (4.128) can be written as $M_{\mathbf{p},\mathbf{k}-\mathbf{p}}(\eta) = \chi_p\chi_{k-p}H^2\eta^2p^2\sin^2\theta$. In the superhorizon limit $|k\eta| << 1$ we have $\chi_p = \frac{\sqrt{-\pi\eta}}{2a}H_\mu^{(1)}(y_p) \sim \frac{iH}{\sqrt{2}}(-\eta)^{-\epsilon}p^{-3/2-\epsilon}$, independent of time on superhorizon scales. Taking into account all this, the time integral in (4.130) yields

$$\int_{\eta_{in}}^0 d\eta' \frac{\mathscr{G}_k^R(\eta,\eta')}{\eta'^3} M_{\mathbf{p},\mathbf{k}-\mathbf{p}}(\eta') \approx \frac{H^4}{4\eta^{1+2\epsilon}}\frac{1}{p^{3/2+\epsilon}}\frac{1}{|\mathbf{k}-\mathbf{p}|^{3/2+\epsilon}}y_p^2\sin^2\theta\left[1-Max\left[\frac{1}{y_p},\frac{1}{y_{|\mathbf{k}-\mathbf{p}|}}\right]^{2+2\epsilon+\delta}\right],$$

and with this, the power spectrum of curvature perturbations reads

$$P_\zeta(\eta) \approx \frac{y_k^3 H^4}{2(8\pi)^2 M_{Pl}^4}\int_0^\infty dy_p\, y_p^{3-2\epsilon}\int_{-1}^{+1} d\cos\theta \frac{1}{y_{|\mathbf{k}-\mathbf{p}|}^{3+2\epsilon}}\sin^4\theta Max\left[\frac{1}{y_p},\frac{1}{y_{|\mathbf{k}-\mathbf{p}|}}\right]^{4+4\epsilon+2\delta} \tag{4.132}$$

where we introduced the Planck mass $M_{pl} \equiv (8\pi G)^{-1/2}$. The following consideration determines the minimum momenta:

$$y_{|\mathbf{k}-\mathbf{p}|}^2 - y_p^2 = y_k^2 - 2y_k y_p \cos\theta = y_k(y_k - 2y_p\cos\theta), \quad y_k, y_p \geq 0\,. \tag{4.133}$$

If $y_k > 2y_p$, then $y_{k-p} > y_p$ for all angles, $-1 < \cos\theta < 1$. For $y_k < 2y_p$ we have to constrain the angle of integration. If $-1 < \cos\theta < 0$, then $y_{k-p} > y_p$ directly. Moreover, this is satisfied also if $0 < \cos\theta < \frac{y_k}{2y_p}$. Then,

$$-1 < \cos\theta < \frac{y_k}{2y_p}, \quad \Longrightarrow \quad y_{k-p} > y_p\,. \tag{4.134}$$

On the other hand,

$$\frac{y_k}{2y_p} < \cos\theta < 1, \quad \Longrightarrow \quad y_{k-p} < y_p\,. \tag{4.135}$$

Therefore, we split the integral (4.132) into three pieces, as follows

$$\begin{aligned} P_\zeta(\eta) \quad &\approx \quad \frac{y_k^3 H^4}{2(8\pi)^2 M_{Pl}^4}\left\{\int_0^{\frac{y_k}{2}} \frac{dy_p}{y_p^{1+6\epsilon+2\delta}} \int_{-1}^{1} \frac{dx}{y_{k-p}^{3+2\epsilon}}(1-x^2)^2\right. \\ &+ \int_{\frac{y_k}{2}}^{\infty} \frac{dy_p}{y_p^{1+6\epsilon+2\delta}} \int_{-1}^{\frac{y_k}{2y_p}} \frac{dx}{y_{k-p}^{3+2\epsilon}}(1-x^2)^2 \\ &+ \left.\int_{\frac{y_k}{2}}^{\infty} dy_p\, y_p^{3-2\epsilon} \int_{\frac{y_k}{2y_p}}^{1} \frac{dx}{y_{k-p}^{7+6\epsilon+2\delta}}(1-x^2)^2\right\}. \end{aligned} \tag{4.136}$$

The integrals can be solved analytically, for example, with Mathematica. The first and third integrals have to be regularized properly to avoid an IR divergence (there are poles at $\mathbf{p} = 0$ and $\mathbf{p} = \mathbf{k}$). Dimensional regularization amounts in taking the substitution $y_p^{1-s} \to y_p^{1-s}$, $(1-x^2)^2 \to (1-x^2)^{2+s/2}$, $y_k^3 \to y_k^{3-s}$, and $H \to H\cdot(H/H_\star)^{s/2}$, where we introduced a parameter $H_\star$ that compensates the dimensions of H (this is called the renormalization constant). To leading order in the slow-roll parameters, the final result is

$$P_\zeta(\eta) \quad \approx \quad \frac{8H^4}{15(8\pi)^2 M_{pl}^4}\left[c + \frac{2n_T}{n_T + n_s - 1}\log\frac{H_k}{H_\star}\right]\left(\frac{k}{k_\star}\right)^{n_s-1}. \tag{4.137}$$

where $n_s - 1 = -4\epsilon - 2\delta$ and $n_T = -2\epsilon$ are the scalar and tensor spectral indices, respectively; and c is a constant. The result is nearly the same as the scalar inflaton spectrum. In addition, there is an implicit logarithmic k-dependence in $H \equiv H_k$ which comes from the contribution of the set of N light scalar fields. The constant term cannot discriminated from the inflaton contribution, because the observational power spectrum is measuring the whole curvature perturbation power spectrum. It is important to remark that the logarithmic contribution comes from the regularization process of the $\mathbf{p}=0$ pole, only. Note: for constant ϵ one has [189] $\delta=-\epsilon$ and consequently $\frac{2n_T}{n_T+n_S-1}=1$. We use this simplified version in the main text.

E. The tensor power spectrum

Tensor modes are determined by the traceless and divergenceless symmetric contribution of the perturbed metric at linear order: $g^T_{\mu\nu}dx^\mu dx^\nu = a^2\left[-d\eta^2+(\delta_{ij}+\gamma_{ij}(t,\mathbf{x}))dx^i dx^j\right]$. In first order perturbation theory the field $\gamma_{ij}(t,\mathbf{x})$ is assumed to propagate in the unperturbed FLRW background. Homogeneity allows to decompose the field in Fourier modes

$$\gamma_{ij}(t,\mathbf{x}) = \sum_{s=\pm2}\int\frac{d^3\mathbf{k}}{(2\pi)^3}e^{i\mathbf{k}\cdot\mathbf{x}}\hat{\gamma}^s_{ij}(t,\mathbf{k}) = \sum_{s=\pm2}\int\frac{d^3\mathbf{k}}{(2\pi)^3}e^{i\mathbf{k}\cdot\mathbf{x}}\hat{\gamma}^s(t,\mathbf{k})e^s_{ij}(\hat{\mathbf{k}}), \quad (4.138)$$

where $s=\pm2$ denote the two possible helicity states of the graviton, $\hat{\gamma}^s(\mathbf{k},t)=\hat{a}^s_{\mathbf{k}}\gamma_k(t)+\hat{a}^{s\,\dagger}_{-\mathbf{k}}\gamma^*_k(t)$ are the quantum modes, and $e^s_{ij}(\hat{\mathbf{k}})$ is the polarization tensor in direction $\hat{\mathbf{k}}$, which must verify the above mentioned properties: it is divergencefree $k_i e^s_{ij}(\hat{\mathbf{k}})=0$, symmetric $e^s_{ij}(\hat{\mathbf{k}})=e^s_{ji}(\hat{\mathbf{k}})$, and traceless $\delta^{ij}e^s_{ij}(\hat{\mathbf{k}})=0$. For a wave propagating in the $\hat{e}_z$-direction the polarization tensor is of the form

$$e^\pm_{ij}(\hat{e}_z)=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & \pm i & 0\\ \pm i & -1 & 0\\ 0 & 0 & 0\end{pmatrix}. \quad (4.139)$$

For a generic direction $\hat{\mathbf{k}}$, the polarization tensor is given by $e^\pm_{ij}(\hat{\mathbf{k}})=S_{ik}(\hat{\mathbf{k}})S_{jl}(\hat{\mathbf{k}})e^\pm_{kl}(\hat{z})$, where $S_{ij}(\hat{\mathbf{k}})$ is the standard 3-dimensional rotation matrix that takes the $\hat{e}_z$ axis into the $\hat{\mathbf{k}}$-direction. For $\hat{\mathbf{k}}=(\sin\theta\cos\phi,\sin\theta\sin\phi,\cos\theta)$ it reads

$$S_{ij}(\hat{\mathbf{k}})=\begin{pmatrix}\cos\phi\cos\theta & -\sin\phi & \sin\theta\cos\phi\\ \sin\phi\cos\theta & \cos\phi & \sin\theta\sin\phi\\ -\sin\theta & 0 & \cos\theta\end{pmatrix}, \quad S\cdot S^T=S^T\cdot S=\mathbb{I}. \quad (4.140)$$

It is easy to verify that $e^\pm_{ij}=(e^\mp_{ij})^*$, $e^\pm_{ij}(-\hat{\mathbf{k}})=e^\mp_{ij}(\hat{\mathbf{k}})$, and $e^\pm_{ij}e^\pm_{ij}=0$, $e^\pm_{ij}e^\mp_{ij}=2$.

In Fourier space, the differential equation that governs the dynamics of the tensor metric fluctuations $\hat{\gamma}_{ij}$ is

$$\hat{\gamma}''_{ij}+2\mathcal{H}\hat{\gamma}'_{ij}+k^2\hat{\gamma}_{ij}=\frac{2}{M^2_{pl}}\Pi^T_{ij}, \quad (4.141)$$

where Π^T_{ij} are the tensor modes of the anisotropic stress. For a massless and minimally-coupled scalar field χ, the total anisotropic stress is calculated in (4.121). By projecting equation (4.141) with the polarization tensor we obtain

$$\hat{\gamma}''_{\pm} + 2\mathcal{H}\hat{\gamma}'_{\pm} + k^2\hat{\gamma}_{\pm} = -\frac{a^2}{2M_{pl}^2}S_T^{\pm}. \tag{4.142}$$

where $S_T^{\pm} = -\frac{2}{a^2}\Pi^T_{ij}e^{\mp}_{ij}$ is the external source for tensor fluctuations, introduced to emphasize the close analogy with the scalar equation (4.104).

Recalling equation (4.121), since $\Pi_{ij} = \Pi^s_{ij} + \Pi^V_{ij} + \Pi^T_{ij}$ such that $e^{\pm}_{ij}\Pi^s_{ij} = e^{\pm}_{ij}\Pi^V_{ij} = 0$, one obtains

$$\begin{aligned}
e^{\pm}_{ij}(\hat{\mathbf{k}})\Pi^T_{ij}(\hat{\mathbf{k}}) &= e^{\pm}_{ij}(\hat{\mathbf{k}})\Pi_{ij}(\hat{\mathbf{k}}) = e^{\pm}_{ij}(\hat{\mathbf{k}})\int\frac{d^3\mathbf{p}}{(2\pi)^3}p_ip_j\,\hat{\chi}_{\mathbf{p}}\hat{\chi}_{\mathbf{k}-\mathbf{p}} \\
&= \int\frac{d^3\mathbf{p}}{(2\pi)^3}p^2e^{\pm}_{33}(\hat{\mathbf{k}})\hat{\chi}_{\mathbf{p}}\hat{\chi}_{\mathbf{k}-\mathbf{p}} \\
&= \int\frac{d^3\mathbf{p}}{(2\pi)^3}p^2S_{3i}(\hat{\mathbf{k}})S_{3j}(\hat{\mathbf{k}})e^{\pm}_{ij}(\hat{e}_z)\hat{\chi}_{\mathbf{p}}\hat{\chi}_{\mathbf{k}-\mathbf{p}} \\
&= \frac{1}{\sqrt{2}}\int\frac{d^3\mathbf{p}}{(2\pi)^3}p^2\sin^2\theta\,\hat{\chi}_{\mathbf{p}}\hat{\chi}_{\mathbf{k}-\mathbf{p}},
\end{aligned}$$

so that

$$S_T^{\pm} = -\sqrt{2}\int\frac{d^3\mathbf{p}}{(2\pi)^3}\frac{p^2}{a^2}\sin^2\theta\,\hat{\chi}_{\mathbf{p}}\hat{\chi}_{\mathbf{k}-\mathbf{p}}. \tag{4.143}$$

The contributions from both tensor polarizations are exactly equal (a consequency of parity invariance), so that an additional factor 2 must be included when considering the total tensor power spectrum, defined by $\langle\hat{\gamma}_{\mathbf{k}}\cdot\hat{\gamma}_{\mathbf{k}'}\rangle = (2\pi)^3P_\gamma\delta^{(3)}(\mathbf{k}+\mathbf{k}')$.

Comparing with the analogous expression for the scalar source, (4.123), we find that the tensor spectrum is simply 8 times that found for the scalar case (4.137),

$$P_\gamma(k) \approx \frac{H^4}{15\pi^2M_{pl}^4}\left[c + \log\frac{H}{H_\star}\right]\left(\frac{k}{k_\star}\right)^{n_T}. \tag{4.144}$$

CHAPTER 5

ELECTROMAGNETIC DUALITY ANOMALY IN CURVED SPACETIMES

The source-free Maxwell action is invariant under electric-magnetic duality rotations in arbitrary spacetimes. This leads to a conserved classical Noether charge. We show that this conservation law is broken at the quantum level in the presence of a classical and dynamical gravitational background field characterized by a nontrivial Chern-Pontryagin invariant, in parallel with the chiral anomaly for massless Dirac fermions. We discuss physical consequences in astrophysics, among which the net polarization of the quantum electromagnetic field is not conserved.

The work presented in this chapter was developed in collaboration with I. Agullo and J. Navarro-Salas [12–14].

5.1 Introduction

Symmetries play a primary role in all areas of physics. They are widely considered as the guiding principle for constructing any physically viable theory, and their connection with conservation laws found by Noether almost one century ago [142] is a cornerstone in understanding all modern fundamental theories.

A particularly interesting example of this is given by Maxwell theory of electrodynamics, whose invariance under Poincare tranformations, which can also be enlarged to the full conformal transformation group, leads to conservation of energy, linear and angular momentum. The theory is also invariant under gauge transformations when the electromagnetic potential is introduced and coupled to matter fields, and implies the conservation of electric charge. But not only that, in the absence of electric charges or currents, and in four dimensions, there is still another symmetry, which perhaps has received much less attention historically, at least from a dynamical point of view. It is a simple exercise to check that both *source-free* Maxwell equations and stress-energy tensor are manifestly invariant, in 4 dimensions, under a so-called duality transformation of the electromagnetic field, that essentially consists of a U(1) rotation between the electric and magnetic fields. Although apparently trivial, this symmetry can have interesting consequences.

It seems that it was Heaviside [115, 116], after reformulating Maxwell equations in the modern language of vector calculous, the first who pointed out the manifest symmetry of source-free Maxwell dynamics under the exchange of the electric and magnetic fields. Then Raichnich [167] introduced the continuous transformation in modern terms. Around 50 years later, Deser and Teitelboim [70] (see also [69]) proved that this transformation is also a symmetry of the standard action by working with the basic dynamical variables, the electromagnetic potential. The validity of this was extended for an arbitrary curved spacetime, and the corresponding conserved Noether charge was identified. It should be stressed that in Minkowski spacetime this charge was previously recognised as the net difference between right- and left-handed circularly polarized radiation intensity [55]. In optics literature this is called optical helicity [39][1], and it corresponds to the V-stokes parameter describing the polarization state of light. Henceforth, besides conservation of energy or momenta, the polarization of radiation is also a constant of motion as long as no electromagnetic sources are present, courtesy of the classical duality symmetry.

A natural question now is whether this symmetry continues to hold in quantum electrodynamics. Note that it is not trivial to promote a classical symmetry to the quantum theory: Noether's theorem states that the continuity equation for a given current is a linear combination of the equations of motion, with the coefficients being the fields themselves, thereby leading to a quadratic operator. Off-shell contributions coming from quantum corrections might spoil the classical invariance, specially if we have an external background field. When this occurs, we speak of a quantum anomaly in the theory.

Historically, the issue of quantum anomalies first appeared in the seminal works by Adler, Bell and Jackiw [10, 44]. They found that the usual chiral symmetry of a massless

[1]see also [45] for a related — but different — notion: the magnetic helicity

Dirac theory, when the fermionic field is interacting with a classical electromagnetic field, breaks down at the quantum level due to the renormalization process, thus yielding the so-called chiral or axial anomaly. Immediately after, Kimura [122] found a similar anomaly when the massless Dirac field is immersed in a classical gravitational background. These discoveries led to an outbreak of interest in anomalies in the QFT and mathematical physics communities, leading to further examples and a connection with the well-known index-theorems in geometric analysis [81, 140]. From a physical point of view they have important implications, for instance they provide an understanding of the pion decay to two photons [44], or a partial understanding of the Standard Model via anomaly cancelation [174]. They have played a major role in string theories too. A decade later of the discovery, an important insight about the origin of quantum anomalies was given by Fujikawa [91, 92] using the path-integral language [119]. He found that, although the classical action remains invariant, it is the measure of the path integral describing the transition amplitude that fails to be invariant under the transformation of interest, and here the reason for the anomaly.

In this chapter we address the question of whether the classical duality symmetry extends to the quantum electromagnetic theory if a dynamical, classical, curved spacetime background is switched on, and in the total absence of electromagnetic sources. To clarify the question in precise terms, one needs to compute the vacuum expectation value of the continuity equation, $\langle \nabla_\mu j_D^\mu \rangle$, where j_D^μ is the Noether current of the duality transformations. To this matter, we shall stress the role played by a duality transformation in Maxwell theory as that of a generalization of a chiral rotation for a massless Dirac field. Then apply standard techniques from quantum field theory in curved backgrounds such as renormalization, geometric analysis (asymptotic expansion of the heat kernel), and Fujikawa's viewpoint using path integrals. In particular we shall show that a duality anomaly arises in the failure of the measure of the path integral to respect the symmetry of the action. Discussions on promising physical implications of all this in astrophysics and connected with gravitational waves are presented at the end of the chapter.

Main results of this chapter were summarized in [13, 14], and a follow-up paper concerning the rest of results that appear in this chapter are planned to be published somewhere.

Notation. We follow the convention $\epsilon^{0123} = 1$ and metric signature $(+,-,-,-)$. More specifically, we follow the $(-,-,-)$ convention of [136]. We always restrict to 4-dimensional spacetimes. Greek indices denote spacetime indices while Minkowski space carries latin indices. We reserve the letters $i, j, k \dots$ for purely spatial objects. Capital indices I, J, or

$\dot{I}, \dot{J}$ refers to internal indices associated to the three-dimensional Lorentz representations. Sum over repeated indices is understood. We assume $c = 1$ but not necessarily $\hbar = 1$. For any issue related with the electromagnetic theory, we follow notation of [120]. Unless otherwise stated, we assume all fields to be smooth and that decay sufficiently fast at infinity.

5.2 Classical theory and electric-magnetic rotations

Lagrangian formalism

In this paper we are concerned with free Maxwell's theory, i.e. electromagnetic fields in the absence of electric charges and currents, formulated on a globally hyperbolic spacetime with metric tensor $g_{\mu\nu}$. The classical theory is described by the action

$$S[A_\mu] = -\frac{1}{4}\int \mathrm{d}^4x\sqrt{-g}\,F^{\mu\nu}F_{\mu\nu} \tag{5.1}$$

where F is a closed two-form (dF = 0) defined in terms of its potential A as $\mathrm{F} = \mathrm{d}A$, or in components, $\mathrm{F}_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu$. Maxwell's equations read $\Box A_\nu - \nabla^\mu\nabla_\nu A_\mu = 0$, where ∇ is the covariant derivative associated with $g_{\mu\nu}$ and $\Box \equiv g^{\mu\nu}\nabla_\mu\nabla_\nu$. When written in terms of the dual tensor $^\star\mathrm{F}$, these equations take the compact form $\mathrm{d}^\star\mathrm{F} = 0$ and, together with $\mathrm{dF} = 0$, make manifest that dynamics is invariant under electric-magnetic rotations

$$\begin{aligned} \mathrm{F} &\longrightarrow \mathrm{F}\cos\theta + {}^\star\mathrm{F}\sin\theta\,, \\ {}^\star\mathrm{F} &\longrightarrow {}^\star\mathrm{F}\cos\theta - \mathrm{F}\sin\theta\,. \end{aligned} \tag{5.2}$$

For $\theta = \pi/2$ one has the more familiar duality transformation $\mathrm{F} \to {}^\star\mathrm{F}$ and ${}^\star\mathrm{F} \to -\mathrm{F}$. If this one-parameter family of transformations are a true symmetry of the action, then Noether's analysis must provide a conserved charge associated to it. We now analyze this problem. Our presentation re-phrases in a manifestly covariant way the results of [70].

The transformation (5.2) is a symmetry of the action if its infinitesimal version ($\delta\mathrm{F} = {}^\star\mathrm{F}\,\delta\theta$, $\delta^\star\mathrm{F} = -\mathrm{F}\,\delta\theta$) leaves the action invariant or, equivalently, changes the Lagrangian density $\mathscr{L} = -\frac{1}{4}\sqrt{-g}F_{\mu\nu}F^{\mu\nu}$ by a total derivative $\delta\mathscr{L} = \sqrt{-g}\nabla_\mu h^\mu$, for some current h^μ. This must be true even off-shell, i.e. when F and *F do not satisfy the equations of motion. In analyzing if this is the case one faces two difficulties. On the one hand, since F is *by construction* a closed two-form (i.e. dF = 0), for the transformation (5.2) to be consistent *F must be also closed, and this amounts to say that *F satisfies the equations of motion. In other words, the transformation (5.2) is only valid on-shell, and a more general transformation has to be considered for this analysis (i.e. something like $\delta \mathrm{F} = \delta\theta\, dZ$, for some 1-form Z; see below).[2] And secondly, since the configuration variables of Maxwell's action are the vector potential A rather than the field F, to apply Noether's techniques we first need to re-write (5.2) in terms of A. A convenient strategy to deal with these two issues is to define a more general transformation, that will agree with electric-magnetic rotations only on-shell, as follows

$$\delta A_\mu = Z_\mu \delta\theta\,, \tag{5.3}$$

where Z_μ is a functional of A_μ implicitly defined by $\mathrm{d}Z = {}^\star\mathrm{F} + \mathrm{G}$, where G is a two-form whose explicit form will not be needed, but that is subject to the following conditions:

1. It is not closed, $\mathrm{dG} \neq 0$. This guarantees the off-shell character ($d^\star\mathrm{F} \neq 0$) of the whole Noether analysis.

2. It vanishes when A_μ satisfies the equations of motion, $\mathrm{G}|_{on-shell} = 0$. This ensures that the usual electromagnetic transformation between the electric and magnetic fields is recovered, $\delta\mathrm{F}|_{on-shell} = {}^\star\mathrm{F}\,\delta\theta$.

3. Fix an arbitrary time-like vector field n^ν. Then G should have zero magnetic part, i.e. $n^{\nu\star}\mathrm{G}_{\mu\nu} = 0$, with ${}^\star\mathrm{G}$ the dual of G. This is motivated by working in the reduced phase space, in which the source-free Gauss law holds, and this way Z can be understood as the "potential" for the electric field (see Appendix A for more details).

Notice that condition 2 also implies that on-shell Z is simply the potential of the dual field ${}^\star\mathrm{F}$, and (5.3) reproduces then the electric-magnetic rotations [take exterior derivative of (5.3) to see this explicitly]. Note also that Z_μ is a non-local functional of A_μ: indeed, as introduced in more detail in Appendix A, we define Z to be the vector "potential" of the electric field, and in the lagrangian formalism $E_a = n^b(\nabla_b A_a - \nabla_a A_b)$ holds. However, as discussed in [70], this is a non-locality in space, and not in time, and therefore it is not an impediment to apply Noether's formalism.

Under the transformation (5.3), we obtain (see Appendix A)

$$\delta\mathscr{L} = -\delta\theta\,\frac{\sqrt{-g}}{2}\,\nabla_\mu\left[A_\nu{}^\star\mathrm{F}^{\mu\nu} - Z_\nu(\mathrm{d}Z)^{\mu\nu}\right] \equiv \sqrt{-g}\,\nabla_\mu h^\mu\,. \tag{5.4}$$

[2] This "difficulty" is singular of the second order formalism. If one uses a first order Lagrangian, or a Hamiltonian formulation, the usual electric-magnetic rotations can be implemented off-shell. This point has been emphasized in [69] and will be manifest in the rest of this paper.

This confirms that electric-magnetic rotations are a symmetry of source-free Maxwell's theory. The density current associated with this symmetry is

$$j^{\mu}_{D} = (-g)^{-1/2}\left(\frac{\partial \mathscr{L}}{\partial \nabla_{\mu} A_{\nu}}\delta A_{\nu}\right) - h^{\mu} = \frac{1}{2}\left[A_{\nu}{}^{\star}\mathrm{F}^{\mu\nu} - Z_{\nu}\,2\,\mathrm{F}^{\mu\nu} - Z_{\nu}({}^{\star}\mathrm{d}Z)^{\mu\nu}\right] \tag{5.5}$$

(we have dropped $\delta\theta$ from the definition of j^{μ}_{D}). This current is gauge-dependent and non-local in space. But this is not a problem, as long as the associated conserved charge is gauge invariant, which is in fact the case. When evaluated on-shell (i.e. when $\mathrm{d}Z = {}^{\star}\mathrm{F}$)

$$j^{\mu}_{D}|_{on-shell} = \frac{1}{2}\left[A_{\nu}{}^{\star}\mathrm{F}^{\mu\nu} - Z_{\nu}\mathrm{F}^{\mu\nu}\right]. \tag{5.6}$$

Now, if we foliate the spacetime using a one parameter family of Cauchy hyper-surfaces Σ_t, the quantity

$$Q_D = \int_{\Sigma_t} \mathrm{d}\Sigma_{\mu} j^{\mu}_{D} = \frac{1}{2}\int_{\Sigma_t} \mathrm{d}\Sigma_3 \left(A_{\mu}B^{\mu} - Z_{\mu}E^{\mu}\right), \tag{5.7}$$

is a conserved charge, in the sense that it is independent of the choice of "leaf" Σ_t. In this expression, $\mathrm{d}\Sigma_3$ is the volume element in Σ_t; $E^{\mu} := n_{\nu}\mathrm{F}^{\nu\mu}$ and $B^{\mu} := n_{\nu}{}^{\star}\mathrm{F}^{\nu\mu}$ are the electric and magnetic parts, respectively, of the electromagnetic tensor field F relative to the foliation Σ_t. The same expression for Q_D is obtained if $j^{\mu}_{D}|_{on-shell}$ is used in place of j^{μ}_{D} in (5.7), and hence the conserved charge is insensitive to the extension of the transformation done above.

In section 5.4 we will re-derive j^{μ}_{D} by using self-dual and anti self-dual variables, and this will make the derivation significantly more transparent. The physical interpretation of Q_D will become also more clear, and we postpone the discussion until then.

The current is conserved on-shell, as it should. Indeed:

$$\begin{aligned}\nabla_{\mu} j^{\mu}_{D} &= \frac{1}{\sqrt{-g}}\left[[\nabla_{\mu}\frac{\partial \mathscr{L}}{\partial \nabla_{\mu} A_{\nu}}]\delta A_{\nu} + \frac{\partial \mathscr{L}}{\partial \nabla_{\mu} A_{\nu}}\nabla_{\mu}\delta A_{\nu}\right] - \nabla_{\mu}h^{\mu} = \frac{1}{\sqrt{-g}}\left[\nabla_{\mu}\frac{\partial \mathscr{L}}{\partial \nabla_{\mu} A_{\nu}}\right]\delta A_{\nu} \\ &= -Z_{\nu}\nabla_{\mu}F^{\mu\nu}\end{aligned} \tag{5.8}$$

so that, by imposing Maxwell equations $\nabla_{\mu}F^{\mu\nu}|_{on-shell} = 0$, we find the desired result.

Ultimately, the question that we want to address in this chapter is whether the symmetry continues to hold at the quantum level. This means that we have to calculate the vacuum expectation value of the divergence of the current and check if it vanishes or not [3]. Notice that this is not a trivial question, since the RHS of (5.8) will be a composite

[3] There are good reasons to expect a non-vanishing value for the right-hand-side of (5.8). Explicit calculations for the local expectation values of $\langle F_{\mu\nu}F^{\mu\nu}\rangle$ [15], $\langle F_{\mu\nu}{}^{\star}F^{\mu\nu}\rangle$ [73, 169] produce in a non-zero result. If the symmetry were preserved quantum-mechanically and the vacuum state were invariant under this transformation, vacuum expectation values of operators that reverse sign under a discrete duality transformation [i.e. $\vec{E} \to \vec{B}$, $\vec{B} \to -\vec{E}$], such as $F_{\mu\nu}F^{\mu\nu} = 2[E_{\mu}E^{\mu} - B_{\mu}B^{\mu}]$, or $\langle F_{\mu\nu}{}^{\star}F^{\mu\nu}\rangle = -4E_{\mu}B^{\mu}$ should vanish.

operator in the quantum theory and this need not vanish. The subtractions required to renormalize expectation values of quadratic operators fail, in general, to respect the classical relations between fields, even if they involve the equations of motion. The main difficulty in this calculation is the fact that it is a non-local functional of A. To overcome this inconvenient we shall approach the question from the first-order or Hamiltonian point of view. Working in phase space allows two additional degrees of freedom per spacetime point to work with, which we can assign to the potential Z to consider it local in space. The cost of doing this, as we shall see, is that the fundamental Poisson brackets will not be local, but this aspect will not be problem at the end of the day.

Hamiltonian formalism

The Hamiltonian formalism provides a complementary approach to the study of the electric-magnetic symmetry, and in this subsection we briefly summarize the derivation of Q_D in this framework. We will restrict here to Minkowski spacetime, since the generalization to curved geometries using the standard vector potential and electric field as canonical coordinates becomes cumbersome. The extension to curved spacetime will become more straightforward once we introduce the self-dual and anti self-dual variable, and we postpone the discussion to section 5.4 and Appendix D.

Given an inertial frame in Minkowski spacetime, Maxwell's Lagrangian (5.1) takes the form

$$L[A,\dot{A}] = \int \mathrm{d}^3x\, \mathscr{L}[A,\dot{A}] = \int \mathrm{d}^3x\, \frac{1}{2}\left[(-\dot{\vec{A}} - \vec{\nabla}A_0)^2 - (\vec{\nabla}\times\vec{A})^2\right], \tag{5.9}$$

where $\vec{\nabla}$ is the usual three-dimensional derivative operator. From this Lagrangian, we see that the canonically conjugate variable of $\vec{A}$ is the negative of the electric field $\frac{\delta L}{\delta \dot{\vec{A}}} = (\dot{\vec{A}} + \vec{\nabla}A_0) = -\vec{E}$, and the conjugate variable π_{A_0} of A_0 vanishes, since the Lagrangian does not involve $\dot{A}_0$. This introduces the constraint $\pi_{A_0} = 0$. Then A_0 is simply a Lagrange multiplier, and from its equation of motion one obtains a constraint, the familiar Gauss' law $\vec{\nabla}\cdot\vec{E} = 0$. Normally one works in the reduced phase space in which $\pi_{A_0} = 0$ always holds and A_0 is an (otherwise arbitrary) fixed function. Then, the phase space can be taken as made of pairs $(\vec{A},\vec{E})$, with a symplectic, or Poisson structure given by $\{A_i(\vec{x}), E^j(\vec{x}')\} = -\delta_i^j\,\delta^{(3)}(\vec{x}-\vec{x}')$. A Legendre transformation gives rise to the Hamiltonian

$$H[A,E] = \int \mathrm{d}^3x\, \frac{1}{2}\left[\vec{E}^2 + (\vec{\nabla}\times\vec{A})^2 - A_0(\vec{\nabla}\cdot\vec{E})\right] \tag{5.10}$$

In Dirac's terminology, $\vec{\nabla}\cdot\vec{E} = 0$ is a first class constraint, and tells us that there is a gauge freedom in the theory, given precisely by the canonical transformations generated by $\vec{\nabla}\cdot\vec{E}$.

Hamilton's equations read

$$\begin{aligned} \dot{\vec{A}} &= \{\vec{A},H\} = -\vec{E} - \vec{\nabla}A_0 \\ \dot{\vec{E}} &= \{\vec{E},H\} = \vec{\nabla}\times(\vec{\nabla}\times\vec{A}) . \end{aligned} \tag{5.11}$$

where $A_0(\vec{x})$ is now interpreted as an arbitrary function without dynamics, and the term proportional to it in the expression for $\dot{\vec{A}}$ corresponds precisely to the gauge flow. These six equations, together with the Gauss constraint, are equivalent to standard Maxwell's equations (once we define $\vec{B}\equiv\vec{\nabla}\times\vec{A}$). To get to the 4 dimensionality per spacetime point of the physical phase space of electrodynamics, one has to fix the gauge.

Electric-magnetic rotations in phase space are given by

$$\delta\vec{E} = (\vec{\nabla}\times\vec{A}) \equiv \vec{B} , \qquad \delta\vec{A} = -(\vec{\nabla}\times)^{-1}\vec{E} \equiv \vec{Z} , \tag{5.12}$$

where $(\vec{\nabla}\times)^{-1}$ is the inverse of the curl; when acting on traverse fields (such as $\vec{E}$) it can be easily computed using the relation $(\vec{\nabla}\times)^{-1} = -\nabla^{-2}\vec{\nabla}\cdot$ The presence of the operator $(\vec{\nabla}\times)^{-1}$ in (5.12) makes evident that we are dealing with a transformation that is non-local in space. Now, the generator of the transformation (5.12) can be easily obtained by computing the symplectic product of $(\vec{A},-\vec{E})$ and $(\delta\vec{A},-\delta\vec{E})$:

$$Q_D = \Omega[(\vec{A},-\vec{E}),(\delta\vec{A},-\delta\vec{E})] = \frac{1}{2}\int d^3x[\vec{A}\cdot\delta\vec{E} - \delta\vec{A}\cdot\vec{E}] = \frac{1}{2}\int d^3x[\vec{A}\cdot\vec{B} - \vec{Z}\cdot\vec{E}] . \tag{5.13}$$

We obtain in this way the same result as in the previous section [compare with equation (5.7)]. Q_D is independent of A_0, and integrating by parts equation (5.13) it is easy to show that only the transverse part of $\vec{A}$ and $\vec{Z}$ contribute to Q_D; hence it is gauge invariant. It is also a straightforward to check that Q_D is indeed the correct generator, since $\delta\vec{A} = \{\vec{A},Q_D\}$ and $\delta\vec{E} = \{\vec{E},Q_D\}$ produce expressions that agree with (5.12). Since the transformation is generated by a charge then the duality transformation (5.12) is canonical in phase space.

To finish, one can now check that $\dot{Q}_D = \{Q_D,H\} = 0$. Therefore, Q_D is a constant of motion, and the canonical transformations it generates are a symmetry of the theory.

5.3 Electrodynamics in terms of self- and anti self-dual variables

Many aspects of the Maxwell's theory *in absence of charges and currents* become more transparent when self- and anti self-dual variables are used (see e.g. [74, 75, 170, 187]).

Several advantages of these variables are well-known and, in particular, they are commonly used in the spinorial formulation of electrodynamics [100, 143, 164]. Our interest here lies mainly in the fact that the duality transformation can be simply described by a chiral transformation using this language, and this will help us in following the clue to Dirac fields and the associated quantum chiral anomaly. For the sake of clarity, we introduce these variables first in Minkowski spacetime, and extend the formalism later to curved geometries.

Minkowski spacetime

The self- and anti self-dual components of of the electromagnetic field are defined as $\vec{H}_\pm \equiv \frac{1}{\sqrt{2}}(\vec{E} \pm i\vec{B})$. We now enumerate the properties and interesting aspects of these complex variables.

1. **Electric-magnetic rotations**

   The duality transformation rule of the electric and magnetic field

$$\begin{aligned} \vec{E} &\longrightarrow \vec{E}\cos\theta + \vec{B}\sin\theta\,, \\ \vec{B} &\longrightarrow \vec{B}\cos\theta - \vec{E}\sin\theta\,, \end{aligned} \tag{5.14}$$

   translates to

$$\vec{H}_\pm \longrightarrow e^{\mp i\theta}\vec{H}_\pm\,. \tag{5.15}$$

   An ordinary duality transformation $\star\vec{E} = \vec{B}$, $\star\vec{B} = -\vec{E}$ corresponds to $\theta = \pi/2$. Then, the operator $i\star$ produces[4] $i\star\vec{H}_\pm = \pm\vec{H}_\pm$. It is for this reason that $\vec{H}_+$ and $\vec{H}_-$ are called the self- and anti self-dual components of the electromagnetic field, respectively.

2. **Lorentz transformations**

   The components of $\vec{E}$ and $\vec{B}$ mix with each other under a Lorentz transformation. For instance, under a boost of velocity v in the x-direction

$$\begin{aligned} \vec{E} = (E_x, E_y, E_z) &\longrightarrow [E_x, \gamma(E_y - vB_z), \gamma(E_z + vB_y)]\,, \\ \vec{B} = (B_x, B_y, B_z) &\longrightarrow [B_x, \gamma(B_y + vE_z), \gamma(B_z - vB_y)]\,, \end{aligned} \tag{5.16}$$

[4]It is common to add the imaginary unit i because in that way the operator $i\star$ has real eigenvalues, and it can be represented by a self-adjoint operator in the quantum theory.

where $\gamma = 1/\sqrt{1-v^2}$. This transformation does not correspond to any irreducible representation of the Lorentz group. However, when $\vec{E}$ and $\vec{B}$ are combined into $\vec{H}_\pm$, it is easy to see that the components of $\vec{H}_+$ and $\vec{H}_-$ no longer mix

$$\vec{H}_\pm = (H^x_\pm, H^y_\pm, H^z_\pm) \longrightarrow [H^x_\pm, \gamma(H^y_\pm \pm i\, v\, H^z_\pm), \gamma(H^z_\pm \mp i\, v\, H^y_\pm)]. \tag{5.17}$$

These are the transformation rules associated to the two irreducible representations of the Lorentz group for fields of spin $s=1$. They are the so-called $(1,0)$ representation for $\vec{H}_+$, and the $(0,1)$ one for $\vec{H}_-$. More generally, for any element of the restricted Lorentz group $SO^+(1,3)$ (rotations + boots), the infinitesimal transformation reads

$$H^J_\pm \to [D(\epsilon_{ab})]_{IJ} H^J_\pm = \left[\delta_{IJ} + \epsilon_{ab}\, {}^\pm\Sigma^{ab}{}_{IJ}\right] H^J_\pm\,, \tag{5.18}$$

where ${}^\pm\Sigma^{ab}{}_{IJ}$ are the generators of the $(1,0)$ and $(0,1)$ representations, and the anti-symmetric matrix $\epsilon_{ab} = \epsilon_{[ab]}$ contains the parameters of the transformation[5]. This makes transparent the fact that electrodynamics describes fields of spin $s=1$, something that is more obscure when working with $\vec{E}$ and $\vec{B}$, the field strength F, or even the vector potential A_μ.

3. **Maxwell's equations**

The equations of motions for $\vec{E}$ and $\vec{B}$,

$$\begin{aligned} \vec{\nabla}\cdot\vec{E} &= 0\,, \qquad & \vec{\nabla}\cdot\vec{B} = 0\,, \\ \vec{\nabla}\times\vec{E} &= -\partial_t\vec{B}\,, \qquad & \vec{\nabla}\times\vec{B} = \partial_t\vec{E}\,. \end{aligned} \tag{5.19}$$

when written in terms of $\vec{H}_\pm$, take the form

$$\vec{\nabla}\cdot\vec{H}_\pm = 0\,, \qquad \vec{\nabla}\times\vec{H}_\pm = \pm i\,\partial_t\vec{H}_\pm\,. \tag{5.20}$$

Notice that, in contrast to $\vec{E}$ and $\vec{B}$, the self and anti self-dual fields are not coupled by the dynamics. The equations for $\vec{H}_-$ and $\vec{H}_+$ are related by complex conjugation.

This is a linear theory, and the space of solutions has vector space structure. It is spanned by positive- and negative-frequency solutions:

$$\vec{H}_\pm(t,\vec{x}) = \int \frac{\mathrm{d}^3k}{(2\pi)^3}\left[h_\pm(\vec{k})\,e^{-i(k\,t-\vec{k}\cdot\vec{x})} + \bar{h}_\mp(\vec{k})\,e^{i(k\,t-\vec{k}\cdot\vec{x})}\right]\hat{\epsilon}_\pm(\vec{k})\,, \tag{5.21}$$

[5]Along this paper we use latin indices $a,b,c,\cdots$ for tensors in (four dimensional) Minkowski space, and Greek indices $\mu,\nu,\alpha,\cdots$ for curved spaces. Upper case latin indices $I,J,K,\cdots$ take values from 1 to 3.

where $k = |\vec{k}|$, and $h_\pm(\vec{k})$ are complex-valued functions that indicate the amplitude of the positive and negative frequency components of a particular solution. The polarization vectors are given by $\hat{e}_\pm(\vec{k}) = \frac{1}{\sqrt{2}}(\hat{e}_1(\vec{k}) \pm i\,\hat{e}_2(\vec{k}))$ where $\hat{e}_1(\vec{k})$ and $\hat{e}_2(\vec{k})$ are any two unit-vectors that, together with $\hat{k}$, form an orthonormal triad of space-like vectors, with orientation defined by $\hat{e}_1(\vec{k}) \times \hat{e}_2(\vec{k}) = +\hat{k}$.

The explicitly form of a generic solution (5.21) helps to understand the relation between self- or anti self-duality and helicity in Minkowski spacetime. By paying attention to the way the electric and magnetic parts (i.e. the real and imaginary parts of $\vec{H}_\pm$, respectively) rotate with respect to the direction of propagation $\hat{k}$ during the course of time, one finds the following relation:

- Positive-frequency Fourier modes $h_\pm(\vec{k})\,e^{-i(kt-\vec{k}\cdot\vec{x})}\,\hat{e}_\pm(\vec{k})$ have positive helicity (that corresponds to left-handed circular polarization) for self-dual fields, and negative helicity for anti self-dual fields.
- For negative-frequency modes $\bar{h}_\mp(\vec{k})\,e^{i(kt-\vec{k}\cdot\vec{x})}\,\hat{e}_\pm(\vec{k})$ the relation is inverse: they have negative helicity (right-handed circular polarization) for self-dual fields, and positive helicity for anti self-dual fields.

We see that duality and helicity are closely related concepts in Minkowski spacetime, although the relation is not trivial; one needs to distinguish between self- and anti self-dual fields *and* positive and negative frequencies. This is the analog of the familiar relation between chirality and helicity for spin 1/2 fermions. In this sense, duality is the chirality of photons.

In more general spacetimes where neither Fourier modes nor the notions of positive and negative frequency are anymore useful, self- or anti self-duality generalizes the concept of helicity, or handedness of electromagnetic waves.

4. **Self- and anti self-dual potentials**

The constraints $\vec{\nabla}\cdot\vec{H}_\pm = 0$ allow us to define the potentials $\vec{A}_\pm$ by:

$$\vec{H}_\pm = \pm i\,\vec{\nabla} \times \vec{A}_\pm. \tag{5.22}$$

It is clear from this definition that the longitudinal part of $\vec{A}_\pm$ contains a gauge ambiguity $\vec{A}_\pm \to \vec{A}_\pm + \vec{\nabla}A^0\pm$, where $A^0_+(t,\vec{x})$ and $A^0_-(t,\vec{x})$ are arbitrary functions, complex conjugate from each other. Note that no time derivatives have been involved in the definition of these potentials.

5. **Maxwell's equations for potentials**

Substituting (5.22) in the field equations (5.20), produces

$$\pm i\vec{\nabla}\times\vec{A}_{\pm} = -\partial_t\vec{A}_{\pm} + \vec{\nabla}A^0_{\pm}\,. \tag{5.23}$$

These equations by themselves are equivalent to Maxwell's equations. It may be surprising at first that Maxwell's theory can be written as first order equations for potentials. The reason comes from the fact that in—and only in—the source-free theory, in addition to the standard potential $\vec{A}$ defined from $\vec{B}=\vec{\nabla}\times\vec{A}$, Gauss's law $\vec{\nabla}\cdot\vec{E}=0$ allows us to define a second potential $\vec{Z}$, such that $\vec{E}\equiv-\vec{\nabla}\times\vec{Z}$. Then, the first order equations

$$\begin{aligned}\dot{\vec{A}} &= \vec{\nabla}\times\vec{Z}+\vec{\nabla}A^0\,,\\ \dot{\vec{Z}} &= -\vec{\nabla}\times\vec{A}+\vec{\nabla}Z^0\,,\end{aligned} \tag{5.24}$$

are equivalent to Maxwell equations (to see this, take curl and use the relation between potentials and fields). Therefore, Maxwell's equations can be written as first order equations for potentials at the expenses of duplicating the number of potentials. The relation between the two sets of potentials is $A^a_{\pm}=\frac{1}{\sqrt{2}}(A^a\pm i\,Z^a)$.

6. **Manifestly Lorentz-covariant equations**

The equations (5.20) and (5.23) for fields and potentials can be re-written in a more compact way as

$$\alpha_I^{ab}\partial_a H^I_+ = 0\,,\qquad \bar{\alpha}_I^{ab}\partial_a A_{+b} = 0\,. \tag{5.25}$$

The equations for H_- and A_- are obtained by complex conjugation. In these expressions α_I^{ab} are three 4×4 matrices, for $I=1,2,3$, and the bar over α_I^{ab} indicates complex conjugation. The components of these matrices in an inertial frame can be identified by comparing these equations with (5.20) and (5.23):

$$\alpha_1^{ab}=\begin{pmatrix}0&1&0&0\\-1&0&0&0\\0&0&0&-i\\0&0&i&0\end{pmatrix},\quad \alpha_2^{ab}=\begin{pmatrix}0&0&1&0\\0&0&0&i\\-1&0&0&0\\0&-i&0&0\end{pmatrix},\quad \alpha_3^{ab}=\begin{pmatrix}0&0&0&1\\0&0&-i&0\\0&i&0&0\\-1&0&0&0\end{pmatrix} \tag{5.26}$$

This matrices are anti-symmetric ($\alpha_I^{ab}=\alpha_I^{[ab]}$) and self-dual ($i\star\alpha_I^{ab}=\alpha_I^{ab}$). As mentioned above, the equations for the potentials can be obtained from the equations for the fields. The reverse is also true. Therefore, either set of equations completely describes the theory.

Equations for the field similar to $\alpha_I^{ab}\partial_a H_+^I = 0$ have been written before in [75, 187], and are also equivalent to Maxwell's equations when written in the language of spinors [170]. For the potentials, we are not aware of a treatment similar to the one presented in this paper. Although electromagnetic potentials have been defined in the spinorial language [143, 164, 170], to the best of our knowledge their generalization to curved spacetimes have not been developed. We do this in the next section.

7. **Relation between $\vec{H}_\pm$ and the field strength F_{ab}**

   From the fields strength F and its dual *F, we define the self and anti self-dual two-forms $\mathrm{F}_\pm = \frac{1}{\sqrt{2}}(\mathrm{F} \pm i\, ^\star\mathrm{F})$, obeying $i^\star\mathrm{F}_\pm^{ab} = \pm\mathrm{F}_\pm^{ab}$. The relation with $\vec{H}_\pm$ is then given by
$$\mathrm{F}_+^{ab} = \alpha_I^{ab} H_+^I, \qquad \mathrm{F}_-^{ab} = \bar{\alpha}_{\dot{I}}^{ab} H_-^{\dot{I}}. \tag{5.27}$$
   These relations imply that one can understand the three α_I^{ab} matrices as a basis (thus complete and orthonormal) of the three-dimensional space of self-dual fields in Minkowski spacetime. Then H_+^I are simply the components of F_+^{ab} in this basis. Similarly, $\bar{\alpha}_I^{ab}$ provides a basis for anti self-dual fields.

   On the other hand, by using the relations (5.27), and the fact that the α_I-matrices are constant in spacetime, so they are transparent to derivatives, the field equations $\alpha_I^{ab}\partial_a H_+^I = 0$ and $\bar{\alpha}_I^{ab}\partial_a H_-^I = 0$ can be written as $\partial_a \mathrm{F}_+^{ab} = 0$ and $\partial_a \mathrm{F}_-^{ab} = 0$, which are equivalent Maxwell's equations in their more standard form $\mathrm{dF} = 0$, $\mathrm{d}^\star\mathrm{F} = 0$.

8. **Properties of the α_I^{ab} matrices**

   To better understand the properties of the α-matrices, and to generalize them to curved spacetimes (next section), it will be convenient to take a slightly more abstract viewpoint and think about the field H_+^I as belonging to and abstract complex three-dimensional vector space V, that is support of a $(1,0)$ irreducible representation of the Lorentz group. This space is isomorphic to the space of self-dual tensors F_+ in Minkowski spacetime, and α_I^{ab} provides the isomorphism. Furthermore, if V is equipped with an Euclidean inner product δ_{IJ}, then α_I^{ab} provides an isometry between these two complex vector spaces.[6] This viewpoint makes clearer the analogy between the α_I^{ab} and the Pauli matrices σ_i^{AB} (recall that σ_i^{AB} provides an isometry between spatial vectors and $SU(2)$ spinors). The

[6] I.e., given any two self-dual tensors $^{(1)}\mathrm{F}_+^{ab}$ and $^{(2)}\mathrm{F}_+^{ab}$, the isomorphism satisfies $^{(1)}\mathrm{F}_+^{ab}\, {}^{(2)}\mathrm{F}_+^{cd}\, \eta_{ac}\eta_{bd} = {}^{(1)}H_+^I\, {}^{(2)}H_+^J\, \delta_{IJ}$, where $^{(i)}\mathrm{F}_+^{ab} = \alpha_I^{ab}\, {}^{(i)}H_+^I$.

following properties of α_I can be thought as the analog of the familiar properties of the Pauli matrices:

- Anti-commutation relations: $\{\alpha_I, \alpha_J\} \equiv \alpha^{a}{}_{bI}\, \alpha^{bc}{}_{J} + \alpha^{a}{}_{bJ}\, \alpha^{bc}{}_{I} = -\delta_{IJ}\, \eta^{ac}$.
- Commutation relations: $[\alpha_I, \alpha_J] \equiv \alpha^{a}{}_{bI}\, \alpha^{bc}{}_{J} - \alpha^{a}{}_{bJ}\, \alpha^{bc}{}_{I} = 2\,{}^{+}\Sigma^{ac}{}_{IJ}$.

If $\vec{H}_+$ is an element of the complex vector space V, then $\vec{H}_-$ is an element of $\bar{V}$, the complex conjugate space. Although naturally isomorphic, these two spaces are different, and we will use doted indices on elements of $\bar{V} \ni H^{\dot{I}}_-$. The properties of $\bar{\alpha}^{ab}_{\dot{I}}$ are obtained by complex conjugating the properties of α^{ab}_I written above. The anti-commutation relations are identical. However, the conjugation changes the commutation relation to $[\bar{\alpha}_{\dot{I}}, \bar{\alpha}_{\dot{J}}] = 2\,{}^{-}\Sigma^{ab}{}_{\dot{I}\dot{J}}$, where now it is the generator of the $(0,1)$ representation of the Lorentz group that enters in the equation. Appendix 5.7 contains a detailed list of other properties of these tensors.

9. **Second order equations for the potentials A_{+a}**

We focus on A_{+a}, since the derivation for A_{-a} can be obtained form it by complex conjugation. The fastest way to obtain the familiar second order differential equation for A_{+a} is to take time derivative of (5.23), use commutativity with the curl, and then use again (5.23) to eliminate the first derivative in time in favor of the curl. The result can then be written in covariant form as $\Box A_{+a} - \partial^b \partial_a A_{+b} = 0$.

But we are more interested in a derivation that can be generalized to curved spacetimes.

With the definitions introduced in Appendix B.3, the relation (5.22) between H^I_+ and A_{+a} can be written as the “curl" $H^I_+ = i\,\epsilon^{Iab}\partial_a A_{+b}$. Then, by using (5.113), when A_{+a} satisfies the equations of motion (5.25), $\bar{\alpha}^{ab}_{\dot{I}}\partial_a A_{+b} = 0$, the relation between H^I_+ and A_{+a} reduces to

$$H^I_+ = \frac{1}{2}\alpha^{abI}\partial_{[a}A_{+b]}\,. \tag{5.28}$$

Substituting in the equation of motion for H^I_+, (5.25), we find

$$\alpha^{ab}_I \partial_a H^I_+ = \alpha^{ab}_I \partial_a \alpha^{cdI}\partial_{[c}A_{+d]} = -\Box A_{+c} + \partial_d \partial^c A_{+d} = 0\,, \tag{5.29}$$

where in the second equality we have used that $\alpha^{ab}_I \alpha^{cdI}$ is a projector on self-dual forms, which acts on $\partial_{[c}A_{+d]}$ as the identity operator when A_{+c} satisfies the equations of motion. We show this way that all solutions of the first order equations for $A_{+c}(t,\vec{x})$ are also solutions of the familiar second order equations. Reciprocally, given the RHS of (5.29) we get the LHS with the relation (5.28). The $b=0$ equation allows to write $H^I_+ = i\,\epsilon^{Iab}\partial_a A_{+b}$ again, and by property (5.113) we get again with the first-order equations (5.25) for the potentials. So all solutions to the familiar second-order equations are also solutions to the first-order ones, and therefore we conclude that both sets of equations are equivalent.

Notice that the equations of motion $\bar{\alpha}^{ab}_{\dot{I}}\partial_a A_{+b} = 0$, imply that the two-form $\partial_{[a}A_{+b]}$ must be self-dual. This is because, on the one hand, the anti-symmetry of $\bar{\alpha}^{ab}_{\dot{I}}$ makes that only the anti-symmetric part of $\partial_a A_{+b}$ contributes to the equations and, on the other, because contraction with $\bar{\alpha}^{ab}_{\dot{I}}$ extracts the anti self-dual component of $\mathrm{d}A_+$. Therefore, when the equations of motion hold, A_{+a} is the potential of the self dual field strength, $\mathrm{F}_+|_{on-shell} = \mathrm{d}A_+$. This latter formula can be explicitly checked by contracting with $\alpha^{\mu\nu}_I$ equation (5.28) and using property (5.95) and (5.92).

10. **Conserved current and charge**

In terms of self- and anti-self dual variables, electric-magnetic rotations take the simple form

$$H^I_\pm(x) \to e^{\mp i\theta} H^I_\pm(x), \qquad A^a_\pm(x) \to e^{\mp i\theta} A^a_\pm(x). \tag{5.30}$$

And the on-shell current (5.6) takes the form

$$j^a_D|_{on-shell} = -\frac{i}{2}\left[H^I_+\,\alpha^{ab}_I\,A_{-b} - H^{\dot{I}}_-\,\bar{\alpha}^{ab}_{\dot{I}}\,A_{+b}\right], \tag{5.31}$$

(note that this is manifestly real). By using the form of the generic solution to the field equations (5.21), we find that the conserved charge

$$Q_D = \int \mathrm{d}^3x\, j^0_D|_{on-shell} = \int \frac{\mathrm{d}^3k}{(2\pi)^3}\frac{1}{k}\mathrm{Re}\left[h^2_+(\vec{k}) - h^2_-(\vec{k})\right], \tag{5.32}$$

is proportional to the difference of intensity between the self- and anti-self dual parts of field or, equivalently, the difference between the right and left circularly polarized components—i.e. the net helicity. (Q_D, therefore, has dimensions of angular momentum.) For this reason this quantity is often called the *optical helicity* of the electro-magnetic field. Note that this is nothing but the V-Stokes parameter used to describe the polarization state of electromagnetic radiation. Note in particular that if $h_\pm(\vec{k}) = \bar{h}_\mp(\vec{k})$, corresponding to a wave solution with the same right and left amplitudes, the above charge yields zero.

11. **Energy-momentum tensor**

Maxwell's energy-momentum tensor, when written in terms of self- and anti self-dual variables takes the form

$$T_{ab} = \frac{1}{2}(\mathrm{F}_{ac}\mathrm{F}_b{}^c + {}^\star\mathrm{F}_{ac}{}^\star\mathrm{F}_b{}^c) = \mathrm{F}_{+ac}\,\mathrm{F}_{-b}{}^c = \alpha_{Iac}\,\bar{\alpha}^{\dot{I}}{}_b{}^c H_+^I H_{-\dot{I}}\,, \tag{5.33}$$

that is manifestly invariant under electric-magnetic rotations. The time-time component provides the energy density

$$T_{00} = \alpha_{I0c}\,\bar{\alpha}^{\dot{I}}{}_0{}^c H_+^I H_{-\dot{I}} = \delta^{\dot{I}}_I H_+^I H_{-\dot{I}} = \frac{1}{2}(E^2 + B^2)\,. \tag{5.34}$$

Curved spacetimes

The generalization to curved spacetimes of the formalism just presented follows the strategy commonly used for Dirac spin 1/2 fields. Namely, one first introduces an orthonormal tetrad field, or Vierbein, in spacetime $e^\mu_a(x)$. This non-coordinate basis is defined by $g_{\mu\nu}(x) = \eta_{ab} e^a_\mu(x) e^b_\nu(x)$, with $\eta_{ab} = diag\{+1,-1,-1,-1\}$. With it, the curved spacetime α_I-matices are obtained from the flat space ones α_I^{ab} by

$$\alpha_I^{\mu\nu}(x) = e^\mu_a(x) e^\nu_b(x)\,\alpha_I^{ab}\,, \tag{5.35}$$

The Minkowski metric η_{ab} is replaced by $g_{\mu\nu}(x)$; η_{ab} is used to raise and lower flat-space indices $a,b,c,\cdots$, $g_{\mu\nu}(x)$ for indices in the tangent space of the spacetime manifold $\mu,\nu,\beta,\cdots$, and δ_{IJ} and $\delta_{\dot{I}\dot{J}}$ for spin 1 indices. The matrices $\alpha_I^{\mu\nu}(x)$ satisfy algebraic properties analog of the ones derived in Minkowski space

$$\{\alpha_I, \alpha_J\} \equiv \alpha^\mu{}_{\nu I}\,\alpha^{\nu\beta}{}_J + \alpha^\mu{}_{\nu J}\,\alpha^{\nu\beta}{}_I = -\delta_{IJ}\,g^{\mu\beta}\,. \tag{5.36}$$

$$[\alpha_I, \alpha_J] \equiv \alpha^\mu{}_{\nu I}\,\alpha^{\nu\beta}{}_J - \alpha^\mu{}_{\nu J}\,\alpha^{\nu\beta}{}_I = 2\,{}^+\Sigma^{\mu\beta}{}_{IJ}\,. \tag{5.37}$$

The extension of the covariant derivative ∇_μ is obtained also using standard arguments (see e.g. chapter 3.1 and Appendix A of [33]). Namely, the action of ∇_μ on indices I of fields $H_+^I \in V$ is uniquely determined by demanding compatibility with the isomorphism $\alpha_I^{\mu\nu}(x)$, $\nabla_\mu \alpha_I^{\beta\nu}(x) = 0$. The result, as one would expect, agrees with the text-book expression for the covariant derivative acting on fields of spin s, particularized to $s = 1$

$$\begin{aligned}\nabla_\mu H_+^I &= \partial_\mu H_+^I - (w_\mu)_{ab}\,{}^+\Sigma^{ab}{}^I{}_J H_+^J\,,\\ \nabla_\mu H_-^{\dot{I}} &= \partial_\mu H_-^{\dot{I}} - (w_\mu)_{ab}\,{}^-\Sigma^{ab}{}^{\dot{I}}{}_{\dot{J}} H_-^{\dot{J}}\,,\end{aligned} \tag{5.38}$$

where $2^\pm\Sigma$ are the generators of the $(1,0)$ and $(0,1)$ representations of the Lorentz algebra introduced in the previous section (by direct calculation one can check: $[2^+\Sigma^{ab}, 2^+\Sigma^{cd}] =$

$\eta^{ac}2[^+\Sigma]^{bd} - \eta^{ad}2[^+\Sigma]^{bc} + \eta^{bd}2[^+\Sigma]^{ac} - \eta^{bc}2[^+\Sigma]^{ad})$, and $(w_\mu)_{ab}$ is the standard one form spin-connection:

$$(w_\mu)^a{}_b = e^a_\alpha \partial_\mu e^\alpha_b + e^a_\alpha e^\beta_b \Gamma^\alpha_{\mu\beta}\,, \tag{5.39}$$

where $\Gamma^\alpha_{\mu\beta}$ are the Christoffel symbols.

With this in hand, the generalization is straightforward:

1. **Maxwell's equations for the fields**

$$\alpha_I^{\mu\nu} \nabla_\mu H^I_+ = 0\,, \quad \bar{\alpha}_{\dot{I}}^{\mu\nu} \nabla_\mu H^{\dot{I}}_- = 0\,. \tag{5.40}$$

Note the similarity with Dirac's equation. The relation between $H_\pm$ and the self and anti self-dual parts of the field strength F is given by $\mathrm{F}_+^{\mu\nu} = \alpha_I^{\mu\nu} H^I_+$ and $\mathrm{F}_-^{\mu\nu} = \bar{\alpha}_{\dot{I}}^{\mu\nu} H^{\dot{I}}_+$. With this, and keeping in mind that $\nabla_\mu \alpha_I^{\beta\nu} = 0$, equations (5.40) become manifestly equivalent to Maxwell's equations in their familiar form $\mathrm{dF}_+ = 0 = \mathrm{dF}_-$.

2. **Potentials $A_{\pm\mu}$**

The self- and anti self-dual potentials satisfy the first order equations:

$$\bar{\alpha}_{\dot{I}}^{\mu\nu} \nabla_\mu A_{+\nu} = 0\,, \qquad \alpha_I^{\mu\nu} \nabla_\mu A_{-\nu} = 0\,. \tag{5.41}$$

These are equivalent to Maxwell equations. The proof follows the same steps as in Minkowski spacetime, though technically more tricky (see Appendix C for details of the derivation).

The relation between $A_{\pm\mu}$ and $H^I_\pm$ (before involving any equation of motion) requires of a foliation of spacetime in spatial Cauchy hyper-surfaces Σ_t. Given that, $A_{+\mu}$ and H^I_+ are related by means of the "curl":

$$H^I_+ = i\,\epsilon^{I\mu\nu} \nabla_\mu A_{+\nu} \tag{5.42}$$

(and similarly for $A_{-\mu}$ and $H^{\dot{I}}_-$) where $\epsilon^{I\mu\nu}$ is defined in Appendix B.3 and denotes a generalized, purely spatial antisymmetric tensor.[7]

As shown in appendix C, one can easily see that if $A_{+\nu}$ is a solution of (5.41), then $A_{\pm\mu}$ are potentials for $\mathrm{F}_\pm$, and they also satisfy the second order equation $\Box A^{+\mu} = \nabla_\nu \nabla^\mu A^{+\nu} = 0$. And reciprocally, if $A_{+\nu}$ is a solution to (5.41), then it is a solution of the second order equations.

[7]Notice that this curl is independent of the connexion ∇_μ, due to the antisymmetry of $\epsilon^{I\mu\nu}$ in μ and ν. It is useful to keep this in mind in manipulating expressions involving $H^I_\pm$ and $A_{\pm\mu}$.

5.4 First order Lagrangian formalism: Dirac-type formulation

The goal of this section is to write a Lagrangian for electrodynamics in terms of self- and anti self-dual variables. The similarity of equations (5.40) and (5.41) with Dirac's equation, together with the drawback found in the non-locality of Z that makes difficult the calculation of the expectation value of (5.8), motivate us to write a first order Lagrangian (i.e. linear in time derivatives) and write it in a form that will make Maxwell's theory manifestly analog to Dirac's theory, where the mathematical structures associated with spin $s = 1/2$ will be replaced by their $s = 1$ analogs. This formulation will become very useful in the study of the electric-magnetic rotations in the quantum theory.

Hamiltonian framework

Given a foliation Σ_t in spacetime, the Hamiltonian of the theory can be easily obtained from the time-time component of the energy-momentum tensor (5.33), after adding the Gauss' constraint

$$H[A_+,A_-] = \int d\Sigma_3 \left[\gamma_I^{\dot{I}} H_+^I H_{-\dot{I}} - H_+^I \gamma_I^\mu \nabla_\mu A_-^0 - H_-^{\dot{I}} \gamma_I^\mu \nabla_\mu A_+^0\right], \tag{5.43}$$

where γ_I^μ is defined in Appendix B.3. It is more convenient to take the potentials A_+ and A_- as coordinates in phase space, since this makes it clear that we are describing the correct number of degrees of freedom.[8] H_+^I and $H_-^{\dot{I}}$ are then understood a shorthands for $i\epsilon^{I\mu\nu}\nabla_\mu A_{+\nu}$ and $-i\epsilon^{\dot{I}\mu\nu}\nabla_\mu A_{-\nu}$, respectively. However, the basic Poisson brackets take a simpler form when expressed in in terms of $H_\pm$

$$\{A_{-\mu}(\vec{x}), -H_+^I(\vec{x}')\} = \delta_\mu^I\, \delta(\vec{x},\vec{x}')\,. \tag{5.44}$$

and similarly for $\{A_{+\mu}(\vec{x}), -H_-^{\dot{I}}(\vec{x}')\}$, where $\delta(\vec{x},\vec{x}') \equiv h^{-1/2}\,\delta^{(3)}(\vec{x}-\vec{x}')$, and h is the determinant of $h_{\mu\nu}$, the pull-back of the spacetime metric $g_{\mu\nu}$ on the space-like hyper-surface Σ_t. The Poisson brackets between A_+ and A_- can be derived form the previous ones, and they are different from zero (although A_+ and A_- do not form a Darboux pair)

From the Hamiltonian we see that $A_\pm^0$ are a Lagrange multiplier. For the spatial components of $A_{+\mu}$, Hamilton's equations read

$$\dot{A}_{\pm I} = \{A_{\pm i}, H\} = H_{+I} + D_I A_\pm^0\,, \tag{5.45}$$

[8] We have 8 real degrees of freedom in $A_{+\mu}$ and $A_{-\mu}$and two constraints. Since these constraints are first class, each removes two degrees of freedom, making the reduced phase space four-dimensional.

where $D_I A_+^0 \equiv \gamma_I^\mu \nabla_\mu A_+^0$ is the spatial derivative of A_+^0. These equation, when written in covariant form, are precisely the equations of motion derived in the previos section, $\bar{\alpha}_{\dot{I}}^{\mu\nu} \nabla_\mu A_{+\nu} = 0$, and $\alpha_I^{\mu\nu} \nabla_\mu A_{-\nu} = 0$.The charge generating electric-magnetic rotations $\delta A_\pm^\mu = \mp i\theta A_\pm^\mu$ can be again obtained from the symplectic product

$$Q_D = \Omega[(A_{+\mu}, A_{-\mu}), (\delta A_{+\mu}, \delta A_{-\mu})] = +\frac{i}{2}\int_{\Sigma_t} d\Sigma_3 \left[A_{+\mu} H_-^{\dot{I}} \delta_{\dot{I}}^\mu - A_{-\mu} H_+^I \delta_I^\mu \right]. \tag{5.46}$$

This result agrees with (5.7) if we write it in terms of the real variable.

Lagrangian framework

We now obtain a first order Lagrangian, by Legendre-transforming the Hamiltonian written in the previous section

$$\begin{aligned} S[A_+, A_-] &= \int d^4x \sqrt{-g} \Big[-\frac{1}{2} \dot{A}_{+\mu} H_-^{\dot{I}} \delta_{\dot{I}}^\mu - \frac{1}{2} \dot{A}_{-\mu} H_+^I \delta_I^\mu - H[A_+, A_-] \Big] \\ &= -\frac{1}{2} \int d\Sigma_4 \left[H_-^{\dot{I}} \bar{\alpha}^{\mu\nu}_{\dot{I}} \nabla_\mu A_\nu^+ + H_+^I \alpha^{\mu\nu}_I \nabla_\mu A_\nu^- \right]. \end{aligned} \tag{5.47}$$

where $d\Sigma_4$ is the four-dimensional volume element, that in a given chart reads $d^4x\sqrt{-g}$. The Lagrangian density in (5.47) differs by a total derivative from the standard Lagrangian in (5.1), (see Appendix D), thus leading to the same dynamics. The passing from Minkowski to curved spacetime is achieved using the minimal coupling prescription, as the result is written covariantly. The independent variables in this action are taken $A_\nu^\pm$, and therefore H_+^I and $H_-^{\dot{I}}$ are understood as shorthands for $i\epsilon^{I\mu\nu} \nabla_\mu A_{+\nu}$ and $-i\epsilon^{\dot{I}\mu\nu} \nabla_\mu A_{-\nu}$, respectively. Note that this action is first order in time derivatives of $A_{\pm\mu}$, and second order in spatial derivatives. Extremizing the action with respect to $A_{+\nu}$ produces the desired equations of motion (the explicit calculation is not direct and is left in the Appendix E)

$$\frac{\delta S}{\delta A_{+\mu}} = 0 \quad \longrightarrow \quad \alpha_I^{\mu\nu} \nabla_\mu A_{-\nu} = 0. \tag{5.48}$$

The complex conjugate equation is obtained from $\frac{\delta S}{\delta A_{-\mu}} = 0$.

For the computations presented in the next section it would be convenient to fix the Lorenz gauge, $\nabla_\mu A_\pm^\mu = 0$. There is a remarkably simple way of incorporating this condition in the action (5.47). All we need to do is to extend the domain of the indices I and $\dot{I}$ from $\{1,2,3\}$ to $\{0,1,2,3\}$, and define $\alpha_0^{\mu\nu} = \bar{\alpha}_0^{\mu\nu} \equiv -g^{\mu\nu}$. This is analog to the familiar extension of the Pauli matrices $\vec{\sigma}$ by adding σ^0 (the identity), which conmutes with all σ^i, $i = 1,2,3$. Algebraic properties of the $\alpha_I^{\mu\nu}$-matrices extended in this way can be consulted in the Appendix B.1

To simplify the notation, we will use the same name for the action, although from now on the index I is understood to run from 0 to 3. The equations of motion still take take the same form

$$\bar{\alpha}^{\mu\nu}_{\ \ \dot{I}}\, \nabla_\mu A^+_\nu = 0\ , \qquad \alpha^{\mu\nu}_{\ \ I}\, \nabla_\mu A^-_\nu = 0\ , \tag{5.49}$$

but they now include the Lorenz condition as the equation for $I = 0$ ($\dot{I} = 0$)

$$g^{\mu\nu}\nabla_\mu A^+_\nu = 0\ , \qquad g^{\mu\nu}\nabla_\mu A^-_\nu = 0\ . \tag{5.50}$$

Note that the variables $H^0_\pm$ in the action have the sole role of acting as Lagrange multipliers to enforce Lorenz's condition. In fact, the equations of motion enforce $H^0_\pm$ to be constant (see Appendix G).

Inspection of the action (5.47) reveals that, contrary to the standard Maxwell's Lagrangian, the Lagrangian density in (5.47) is manifestly invariant, $\delta\mathscr{L} = 0$, under electric-magnetic rotations $A^\mu_\pm \to e^{\mp i\theta} A^\mu_\pm$. It is straightforward to derive the Noether's current (see Appendix F)

$$j_D|_{on-shell} = (-g)^{-1/2}\left(\frac{\delta\mathscr{L}}{\delta\nabla_\mu A_{+\nu}}\delta A_{+\nu} + \frac{\delta\mathscr{L}}{\delta\nabla_\mu A_{-\nu}}\delta A_{-\nu}\right)|_{on-shell} = \frac{i}{2}\left[H^{\dot{I}}_-\,\bar{\alpha}^{\mu\nu}_{\ \ \dot{I}}A^+_\nu - H^I_+\,\alpha^{\mu\nu}_{\ \ I}A^-_\nu\right]. \tag{5.51}$$

Using the relation between self- and anti self-dual variables, and ordinary variables A_μ and $\mathrm{F}^{\mu\nu}$, it is straightforward to check that this expression agrees with $j^\mu_D|_{on-shell}$ obtained in section 5.2, equation (5.6). The agreement is in fact off-shell. This is checked in Appendix F.

Dirac-type Lagrangian

The goal of this section is to re-write the action (5.47) in a more convenient form. The new form will make the theory formally similar to Dirac's theory of spin 1/2 fermions and will facilitate the computations in the next sections, as well as the comparison with results derived for fermions.

Integrating by parts (5.47)

$$S = -\frac{1}{4}\int d^4x\sqrt{-g}\left[H^{\dot{I}}_-\,\bar{\alpha}^{\mu\nu}_{\ \ \dot{I}}\nabla_\mu A^+_\nu - A^+_\nu\bar{\alpha}^{\mu\nu}_{\ \ \dot{I}}\nabla_\mu H^{\dot{I}}_- + H^I_+\,\alpha^{\mu\nu}_{\ \ I}\nabla_\mu A^-_\nu - A^-_\nu\,\alpha^{\mu\nu}_{\ \ I}\nabla_\mu H^I_+\right]\ . \tag{5.52}$$

and this result can be written as

$$S[A_+, A_-] = -\frac{1}{4}\int \mathrm{d}\Sigma_4\, \bar{\Psi}\, i\beta^\mu\nabla_\mu\Psi \tag{5.53}$$

where we have defined (see Appendix H)

$$\Psi = \begin{pmatrix} A^+ \\ H_+ \\ A^- \\ H_- \end{pmatrix}, \quad \bar{\Psi} = (A^+, H_+, A^-, H_-), \quad \beta^\mu = i \begin{pmatrix} 0 & 0 & 0 & \bar{\alpha}^\mu \\ 0 & 0 & -\alpha^\mu & 0 \\ 0 & \alpha^\mu & 0 & 0 \\ -\bar{\alpha}^\mu & 0 & 0 & 0 \end{pmatrix}. \tag{5.54}$$

Remark: It is convenient to include an (arbitrary) parameter ℓ^{-1} with dimensions of inverse of length in the definition of Ψ and $\bar{\Psi}$ multiplying $A^\pm$ and compensate it by multiplying the RHS of (5.53) by ℓ. The action, and the physical predictions obtained from it remain obviously invariant, but the introduction of ℓ will make all the components of Ψ and $\bar{\Psi}$ to have the same dimensions (namely, $\sqrt{E/L^3}$). To simplify the notation, we will not write ℓ explicitly, but it should be taken into account in counting the dimensions of expressions containing Ψ and $\bar{\Psi}$.

The exact position of the indices in $A_\pm$, $H_\pm$, α^μ and $\bar{\alpha}^\mu$ can be easily obtained by comparing (5.52), (5.53) and (5.54). We have omitted them in the main body of this paper to simplify the notation, and further details can be found in Appendix H. Equation (5.53) is formally analog to the action of a Majorana 4-spinor, whose lower two components are complex conjugate from the upper ones.

From the algebraic properties of the extended α-matrices, (5.132) and (5.134) , we derive the commutation and anti-commutation properties of β^μ

$$\{\beta^\mu, \beta^\nu\} = 2g^{\mu\nu}\mathbb{I}, \qquad [\beta^\nu, \beta^\nu] = -4 \begin{pmatrix} {}^+\Sigma^{\mu\nu} & 0 & 0 & 0 \\ 0 & {}^+\Sigma^{\mu\nu} & 0 & 0 \\ 0 & 0 & {}^-\Sigma^{\mu\nu} & 0 \\ 0 & 0 & 0 & {}^-\Sigma^{\mu\nu} \end{pmatrix} \tag{5.55}$$

The anti-commutations relations tell us that the matrices β^μ provide a (16×16) representations of the Clifford algebra, while the commutation relations indicate that the sub-algebra formed by the quadratic elements $\beta^{[\mu}\beta^{\nu]}$ contains the $(1,0)$ and $(0,1)$ irreducible representations of the Lorentz group. These are the spin 1 analog of the properties of the Dirac γ^μ matrices (in the Weyl representation). From the properties of the α-matrices, we also have that $\nabla_\nu \beta^\mu(x) = 0$.

We now introduce the "chiral" matrix

$$\beta_5 = \frac{i}{4!}\epsilon_{\alpha\beta\gamma\delta}\beta^\alpha\beta^\beta\beta^\gamma\beta^\delta = \begin{pmatrix} -\mathbb{I} & 0 & 0 & 0 \\ 0 & -\mathbb{I} & 0 & 0 \\ 0 & 0 & \mathbb{I} & 0 \\ 0 & 0 & 0 & \mathbb{I} \end{pmatrix} \tag{5.56}$$

Some properties can be immediately checked out:

$$\{\beta^\mu, \beta_5\} = 0\,, \qquad \beta_5^2 = \mathbb{I}\,, \qquad (\beta_5)^\dagger = \beta_5\;. \tag{5.57}$$

Additional properties of these matrices can be found in Appendix H.

Although the basic variables in the action are the potentials $A_{\pm\mu}$, at the practical level one can work by considering Ψ and $\bar{\Psi}$ as independent fields—note that this is the same as one does when working with Majorana spinors. The equation of motion take the form

$$\frac{\delta S}{\delta\bar{\Psi}} = 0 \quad \longrightarrow \quad i\beta^\mu\nabla_\mu\Psi = 0\,. \tag{5.58}$$

They contain four equations, one for each of the four components of Ψ. The upper two are equations $\bar{\alpha}^{\mu\nu}_{\;\;\bar{I}}\,\nabla_\mu A^+_\nu = 0$ and $\alpha^{\mu\nu}_I\nabla_\mu H^I_+ = 0$. As discussed in previous sections, these two sets of equations are equivalent to each other, once the relations $H^I_+ = i\epsilon^{I\mu\nu}\nabla_\mu A_{+\nu}$ are taken into account. The lower two components in (5.58) produce equations that are complex conjugated from the previous ones.

Second order equations can be calculated

$$(-i\beta^\alpha\nabla_\alpha)\,i\beta^\mu\nabla_\mu\Psi = (\beta^{(\alpha}\beta^{\mu)} + \beta^{[\alpha}\beta^{\mu]})\nabla_\alpha\nabla_\mu\Psi = (\Box + \mathscr{Q})\Psi = 0 \tag{5.59}$$

where

$$\mathscr{Q}\,\Psi \equiv \frac{1}{2}\beta^{[\alpha}\beta^{\mu]}W_{\alpha\mu}\Psi \tag{5.60}$$

with

$$W_{\alpha\mu}\Psi \equiv [\nabla_\alpha, \nabla_\mu]\Psi = \frac{1}{2}R_{\alpha\mu\sigma\rho}\begin{pmatrix} \Sigma^{\sigma\rho} & 0 & 0 & 0 \\ 0 & 2\,{}^{+}\Sigma^{\sigma\rho} & 0 & 0 \\ 0 & 0 & \Sigma^{\sigma\rho} & 0 \\ 0 & 0 & 0 & 2\,{}^{-}\Sigma^{\sigma\rho} \end{pmatrix}\Psi\,. \tag{5.61}$$

where $\Sigma^{\sigma\rho} = {}^{+}\Sigma^{\sigma\rho} + {}^{-}\Sigma^{\sigma\rho}$.

Looking at the expression for $W_{\alpha\mu}\Psi$ we see that it contains real terms, $R_{\alpha\mu\sigma\rho}\Sigma^{\sigma\rho}$, as well as complex ones, $R_{\alpha\mu\sigma\rho}\,{}^{\pm}\Sigma^{\sigma\rho}$. The real terms come from the action of covariant derivatives on $A^\mu_\pm$. Since $A^\mu_\pm$ is a vector in spacetime, the covariant derivatives of them includes a connexion associated to the (1/2,1/2) (real) representation of the Lorentz group—although this does not mean that they transform according to the (1/2,1/2) representation; the do only up to a gauge transformation [187, 190]. The complex terms in $W_{\mu\nu}\Psi$ originate from the (1,0) and (0,1) representations, to which $H_\pm$ are associated with.

The canonical Poisson brackets of the classical theory read

$$\{\Psi_A(t,\vec{x}),\Pi^B_\Psi(t,\vec{x}')\} = \delta^B_A\,\delta(\vec{x},\vec{x}') \tag{5.62}$$

where $\Pi_\Psi \equiv \frac{\delta\mathscr{L}}{\delta\dot{\Psi}} = -\frac{i}{4}\bar{\Psi}\beta^0$. These are promoted to *commutation relations* in the quantum theory. If anti-commutators were rather used to quantize the theory, one would find the quantum propagator to violate causuality, as expected from the spin-statistics theorem. Therefore, in spite of the fermion-like appearance of the formulation used in this section, we are describing a theory bosons, with a symmetric Hilbert space.

Axial current

We now describe how electric-magnetic rotations and their associated conservation law looks like in the language introduced in this section. By using the chiral matrix β_5, the transformation reads

$$\Psi \to e^{i\theta\beta_5}\Psi\,, \qquad \bar{\Psi} \to \bar{\Psi}e^{i\theta\beta_5} \tag{5.63}$$

Notice that they have the same form as chiral transformations for fermions. Looking at the form of β_5 in equation (5.56), it is clear that the upper two components of Ψ, i.e. (A_+,H_+), represent the self-dual, or positive chirality part of the field, while the lower two components (A_-,H_-) contain the anti self-dual, or the negative chiral part. As discussed in previous sections, the Lagrangian density is manifestly invariant under these transformation, and in terms of Ψ the conserved current reads

$$j^\mu_D = \frac{1}{4}\bar{\Psi}\beta^\mu\beta_5\Psi\ . \tag{5.64}$$

The associated Noether charge is

$$Q_D = \int_\Sigma \mathrm{d}\Sigma_\mu j^\mu_D = \frac{1}{4}\int \mathrm{d}\Sigma_3\,\bar{\Psi}\beta^0\beta_5\Psi \tag{5.65}$$

This is equivalent to the expressions for Q_D we have written in previous sections [equations (5.7) and (5.46)] when one expands as in (5.54).

5.5 The quantum anomaly

We present two ways to calculate the anomaly. The first one relies on the familiar subtraction procedure, in which the UV divergences are identified and subtracted in a covariant and self-consistent way. The second one is based on the path integral and Fujikawa's interpretation on the emergence of anomalies. In this approach an anomaly arises because, despite the action remains symmetric under the classical transformation, the measure of the integral fails to be invariant.

5.5.1 Direct computation

Both j^μ_D and $\nabla_\mu j^\mu_D$ are operators quadratic in fields, and therefore the computation of their expectation values must include renormalization subtractions to eliminate the potential divergences and to produce a finite, physically reasonable quantity.[9] The regularization of UV divergences can be achieved by subtracting the short-distance behavior of Green functions $S(x,y)$. The DeWitt-Schwinger scheme is a representative method of this. In this framework the quadratic operator of interest is evaluated at different spacetime points and one subtracts the asymptotic short-distance behavior of the bare quantity by expanding it in powers of derivatives of the metric.

$$\langle \nabla_\mu j^\mu_D \rangle_{ren} = \langle \nabla_\mu j^\mu_D \rangle_{formal} - \langle \nabla_\mu j^\mu_D \rangle_{Ad(4)} \,. \tag{5.66}$$

In this expression, $\langle \nabla_\mu j^\mu_D \rangle_{Ad(4)}$ are the terms of the asymptotic expansion of $\langle \nabla_\mu j^\mu_D \rangle_{formal}$ up to four derivatives of the metric. Alternatively to point-splitting, one can keep the points coincident and introduce a parameter $m > 0$ in the theory — that will serve as a regulator, and will be sent to zero at the end of the calculation—by replacing the wave equation (5.58) by $(D-m)\Psi = 0$, with $D \equiv i\beta^\mu \nabla_\mu$. Then

$$\nabla_\mu j^\mu_D = \lim_{m\to 0} -\frac{i}{4}\bar\Psi \overleftarrow{D}\,\beta_5\,\Psi + \frac{i}{4}\bar\Psi \beta_5 \vec{D}\Psi = \lim_{m\to 0}\frac{i}{2}\,m\,\bar\Psi\beta_5\Psi = \lim_{m\to 0}\frac{i}{2}\,m\,\mathrm{Tr}[\beta_5\Psi\bar\Psi]\,, \tag{5.67}$$

where we have used $\{\beta^\mu, \beta_5\} = 0$ in the first equality, and $\bar\Psi^A (\beta_5)_A^{\ B}\,\Psi_B = (\beta_5)_A^{\ B}\,\Psi_B \bar\Psi^A = \mathrm{Tr}[(\beta_5)_A^{\ B}\,\Psi_B\bar\Psi^C]$ in the second one. If we now make a choice of vacuum state $|0\rangle$: $\langle 0|\,\Psi_m(x)\bar\Psi_m(y)\,|0\rangle = S(x,y,m)$ and we obtain

$$\langle \nabla_\mu j^\mu_D \rangle_{formal} = \lim_{m\to 0}\frac{i}{2}\,m\,\mathrm{Tr}\Big[\beta_5\,S(x,x,m)\Big]\,. \tag{5.68}$$

[9]Recall that one must include all the subtractions that would produce divergences for generic values of the parameters of the theory—mass, coupling to the curvature, etc.—even if the concrete value of these parameters that one is interested in makes the divergences to disappear.

and

$$\langle \nabla_\mu j^\mu_D \rangle_{ren} = \lim_{m\to 0} \frac{i}{2} m \,\mathrm{Tr}\Big[\beta_5 \Big(S(x,x,m) - S(x,x,m)_{Ad(4)}\Big)\Big]. \tag{5.69}$$

In the last expression, $S(x,x')$ contains the information about the vacuum state, while $S(x,x')_{Ad(4)}$ removes the universal ultra-violet divergences. It turns out that, because the operator D is not self-adjoint[10], and because the heat kernel expansion is well-known for second order differential equations, it is convenient to work with the green function associated with the latter. By writing $S(x,y,m) = (D+m)G(x,y,m)$ one gets that $(D^2+m^2)G(x,y,m) = \delta^{(4)}(x,y)$. The asymptotic expansion of $G(x,y,m)$ is now well-known. It is given by[11]

$$G(x,x)_{Ad(4)} = \frac{\hbar}{16\pi^2} \sum_{k=0}^{2} \int_0^\infty d\tau \, e^{-i\tau m^2} (i\tau)^{(k-2)} E_k(x) \tag{5.70}$$

where the functions $E_k(x)$ are local, geometric quantities, built from the metric and its first $2k^{th}$ derivatives. Their specific values are known [103, 157]. For manifolds without boundary they read:

$$\begin{aligned}
E_0(x) &= \mathbb{I} \\
E_1(x) &= \frac{1}{6} R\,\mathbb{I} - \mathcal{Q} \\
E_2(x) &= \left[\frac{1}{72}R^2 - \frac{1}{180} R_{\mu\nu}R^{\mu\nu} + \frac{1}{180} R_{\alpha\beta\mu\nu}R^{\alpha\beta\mu\nu} - \frac{1}{30}\Box R\right]\mathbb{I} \\
&+ \frac{1}{12} W_{\mu\nu}W^{\mu\nu} + \frac{1}{2}\mathcal{Q}^2 - \frac{1}{6} R\mathcal{Q} + \frac{1}{6}\Box\mathcal{Q},
\end{aligned}$$

where the expression for $W_{\mu\nu} \equiv [\nabla_\mu, \nabla_\nu]$ and $\mathcal{Q}(x)$ were given in (5.61) and (5.60). R, $R_{\mu\nu}$, and $R_{\alpha\beta\mu\nu}$ are the Ricci scalar, Ricci tensor, and Riemann curvature tensor, respectively.

It turns out that, because of the symmetry of the classical action, the contribution of $S(x,x')$ to (5.69) vanishes. Therefore, $\langle \nabla_\mu j^\mu_D \rangle_{ren}$ arises entirely from the subtraction terms, $S(x,x')_{Ad(4)}$.This implies that $\langle \nabla_\mu j^\mu_D \rangle_{ren}$ is *independent of the choice of vacuum*. Once the limit $m \to 0$ is taken the fourth order contribution will provide a residual finite term, resulting thus in the anomalous conservation of the current. Notice that the same occurs in the calculation of other anomalies, such as the fermionic chiral anomaly or the trace anomaly.

[10]Precisely, this is the reason why the anomaly arises from a mathematically point of view: D and its adjoint $D^\dagger$ have different zero-eigenvalue multiplicity and the analitical index $\dim\ker D - \dim\ker D^\dagger \neq 0$. The Atiyah-Singer index theorems tell us then that the result will depend on the background topology.

[11] This expression for $G(x,x')_{Ad(4)}$ is obtained by writing $G(x,x')_{Ad}$ in terms of its heat kernel $K(\tau,x,x')$, $G(x,x')_{Ad} = i\hbar\int_0^\infty d\tau\, e^{-im^2\tau} K(\tau,x,x')$, and then using the asymptotic expansion $K(\tau,x,x) \sim \frac{i}{16\pi^2}\sum_{k=0}^\infty (i\tau)^{k-2} E_k(x)$ for $\tau \to 0$. See e.g. [157] for further details.

Taking into account that (for details see Appendix I)

$$\begin{aligned}\mathrm{Tr}[\beta_5 E_0] &= 0, \qquad \mathrm{Tr}[\beta_5\beta^\mu E_0]=0\\ \mathrm{Tr}[\beta_5 E_1] &= 0, \qquad \mathrm{Tr}[\beta_5\beta^\mu E_1]=0\\ \mathrm{Tr}[\beta_5 E_2] &= \frac{1}{12}\mathrm{Tr}[\beta_5 W_{\mu\nu}W^{\mu\nu}] = i\,\frac{1}{3}R_{\alpha\beta\mu\nu}\,{}^\star R^{\alpha\beta\mu\nu}\\ \mathrm{Tr}[\beta_5\beta^\mu E_2] &= 0\end{aligned} \tag{5.71}$$

where ${}^\star R^{\alpha\beta\mu\nu} = \frac{1}{\sqrt{-g}\,2!}\epsilon^{\alpha\beta\sigma\rho}R_{\sigma\rho}{}^{\mu\nu}$ is the dual of the Riemann tensor, equation (5.69) produces:

$$\langle \nabla_\mu j^\mu_D\rangle_{ren} = -\frac{\hbar}{96\pi^2}R_{\alpha\beta\mu\nu}\,{}^\star R^{\alpha\beta\mu\nu}\,. \tag{5.72}$$

A few comments are in order now:

1. This result reveals that quantum fluctuations spoil the conservation of the axial current j^μ_D, and break the classical symmetry under electric-magnetic (or chiral) transformations.

2. The pseudo-scalar $R_{\alpha\beta\mu\nu}\,{}^\star R^{\alpha\beta\mu\nu}$ is known as the Chern-Pontryagin density (its integral across the entire spacetime manifold is the Chern-Pontryagin topological invariant).

3. Only the complex elements in the diagonal of $W_{\mu\nu}$ contribute to (5.71) (see (5.61) and the discussion below that equation). This means that the anomalous non-conservation is linked to the $(1,0)$ and $(0,1)$ representations of the Lorentz group.

4. It is important to notice the parallelism with the chiral anomaly for spin 1/2 fermions. The computations in that case would de identical, except that in (5.71) one would have to used the commutator of covariant derivatives $W_{\mu\nu} = [\nabla_\mu, \nabla_\nu]$ associated with spin 1/2 fields, rather than spin 1. That would change only the numerical coefficient in (5.71) and (5.72).

5.5.2 Path integral formalism

The functional integral for the theory under consideration is[12]

[12]As usual, the inclusion of the Lorentz gauge introduces two ghost scalar fields. These fields do contribute to certain observables, such as the trace anomaly. However, one can check explicitly that they do not affect the computation os $\nabla_\mu j^\mu_D$. It is for this reason that we have not written their contribution to the path integral.

$$Z = \int D\bar{\Psi} D\Psi \, e^{i/\hbar S[\Psi,\bar{\Psi}]} \tag{5.73}$$

The strategy is the following. The generating functional Z is invariant under the replacement $(\Psi,\bar{\Psi}) \to (\Psi' = e^{+i\beta_5\theta}\Psi, \bar{\Psi}' = \bar{\Psi}e^{+i\beta_5\theta})$, since this plays the role of a mere change of variables (it is a canonical transformation in phase space), and the path integral must remain invariant under such a change. Noether's theorem in its second version—in which one considers the parameter of the transformation $\theta(x)$ to be a spacetime function of compact support—, tell us that $\delta S = -\int d^4x\sqrt{-g}\,\theta(x)\nabla_\mu j^\mu_D$. Furthermore, the integral measure $D\bar{\Psi}D\Psi$ could change by a non-trivial Jacobian, $D\bar{\Psi}D\Psi \to J D\bar{\Psi}'D\Psi'$. Then, invariance of Z implies that, quantum mechanically, $J\cdot e^{-i/\hbar\int d^4x\sqrt{-g}\theta(x)\nabla_\mu j^\mu_D}$ must be equal to one. From this we see that quantum anomalies will appear for those classical symmetries that do not leave the measure of the path integral invariant. The value of $\langle\nabla_\mu j^\mu_D\rangle$ can be then determined by J. The goal of this section is to compute these quantities.

The Jacobian J can be determined by using standard functional analysis techniques applied to the wave operator $D := i\beta^\mu\nabla_\mu$. Consider the space of square-integrable complex fields $\Psi(x)$ with respect to the product $\langle\Psi_1,\Psi_2\rangle = \frac{\alpha}{4}\int d^4x\sqrt{-g}\,\bar{\Psi}_1\beta\Psi_2$, where the operator β is an involution defined in Appendix H and α is a parameter with dimensions of action[13]. In terms of the original variables $A_\pm$ and $H_\pm$, the norm of $\Psi(x)$ reads $\langle\Psi,\Psi\rangle = \frac{\alpha}{4}\int d^4x\sqrt{-g}[2|A_+|^2+2|H_+|^2]\geq 0$. Note also that the action can be expressed as $S[\Psi,\bar{\Psi}] = -\langle\Psi,\beta D\Psi\rangle$.

The operator $D^\dagger D$ is self-adjoint with respect to the product $\langle\Psi_1,\Psi_2\rangle$. The self-adjointness guarantees the existence of an orthonormal basis $\{\Psi_n\}_{n\in\mathbb{N}}$ made of eigenstates, $D^\dagger D\Psi_n = \lambda_n^2\Psi_n$. We will denote by a_n the components of a vector Ψ in this basis. An electric-magnetic rotation $\Psi\to\Psi' = e^{+i\theta\beta_5}\Psi$ can be now expressed as a change of the components $a_n \to a'_n = \sum_m C_{nm}a_m$, yielding $C_{nm} = \langle\Psi_n, \beta e^{+i\theta\beta_5}\Psi_m\rangle$. With this, the Jacobian of the transformation reads

$$D\bar{\Psi}D\Psi \to J D\bar{\Psi}'D\Psi', \quad \text{with} \quad J = (det C)^2 = 1 + 2i\mathrm{Tr}\sum_{n\in\mathbb{N}}\langle\Psi_n, \beta\beta_5\theta\Psi_n\rangle + O(\theta^2) \tag{5.74}$$

Now, the invariance of the path integral implies that, quantum mechanically

$$\langle\nabla_\mu j^\mu_D\rangle_{ren} = 2\hbar\frac{\alpha}{4}\sum_{n=0}^{\infty}\bar{\Psi}_n\beta_5\Psi_n\,. \tag{5.75}$$

[13]It is introduced in order to make the product dimensionless and, although $\alpha = \hbar$ would be a natural choice, we leave it unspecified to make manifest that physical observables are independent of it; it cancels out in intermediate steps.

To evaluate this expression we use again the heat kernel approach. The heat kernel of the equation, $D^\dagger D\Psi = 0$, is [157]

$$K(\tau,x,x') = \alpha \sum_{n=0}^{\infty} e^{-i\tau\lambda_N^2}\Psi_n(x)\bar{\Psi}_n(x') \tag{5.76}$$

Then

$$\langle\nabla_\mu j_D^\mu\rangle_{ren} = \frac{1}{2}\hbar\lim_{\tau\to 0}\mathrm{Tr}[\beta_5 K(\tau,x,x')] = i\frac{\hbar}{32\pi^2}\mathrm{Tr}[\beta_5 E_2] = -\frac{\hbar}{96\pi^2}R_{\alpha\beta\mu\nu}{}^\star R^{\alpha\beta\mu\nu}\ . \tag{5.77}$$

where in the second equality we have used the expansion of $K(\tau,x,x')$ for $\tau\to 0$, written in Footnote 11, and in the last equality we have used (5.71). For details of the calculation see Appendix I.

Remark: Recall that the path integral produces transition amplitudes for time-ordered products of operators between the “in" and “out” vacuum. However, since the result for $\langle\nabla_\mu j_D^\mu\rangle_{ren}$ comes entirely from the asymptotic terms in the heat kernel, those depend on the background and are universal for any vacuum state. Therefore, the result (5.77) agrees with the *expectation value* of $\nabla_\mu j_D^\mu$ in any vacuum state.

5.6 Conclusions and future prospects

The above result (5.77) implies that the charge Q_D, classically associated with the duality symmetry of the source-free Maxwell action, is no longer conserved in the quantum theory in a general spacetime.

It is interesting now to estimate in more precise terms which physical implications this anomaly could provide. Since in flat space Q_D represents the difference in number between photons of opposite helicity [55], this result could perhaps be interpreted as a nonconservation of the V-stokes parameter of the quantum electromagnetic field in curved spacetimes. In order to figure out quantitatively the significancy of this, one needs to study the RHS of (5.77). Note first that the Pontryagin density invariant only depends on the vacuum geometry of the background (even if $R_{ab}\neq 0$): $R_{abcd}{}^\star R^{abcd} = C_{abcd}{}^\star C^{abcd} = 16E_{ab}B^{ab}$, where C_{abcd}, E_{ab}, B_{ab} are respectively the Weyl tensor, and its Electric and Magnetic parts. The change in Q_D between two different spatial Cauchy surfaces could be stated as

$$\Delta Q_D = \frac{2\hbar}{3\pi^2}\int_{t_1}^{t_2}\int_\Sigma d^4x\sqrt{-g}\,E_{\mu\nu}B^{\mu\nu}\ . \tag{5.78}$$

An analog of this effect arises for fermions in the creation of pairs from the vacuum by strong electric fields. In this situation the presence of a magnetic field would induce a non-zero net chirality on the particles created, as predicted by the chiral anomaly $\langle \partial_a j^a_{1/2} \rangle \sim E_a B^a$, [141, 193]. Likewise, apart from gravitational tidal forces produced by $E_{\mu\nu}$, frame dragging effects, described by $B_{\mu\nu}$, are necessary to induce net polarization on the created photons.

To illustrate this, consider the process of collapse of a neutron star into a Kerr black hole. In the vicinity of the region where the event horizon will form, the gravitational field is strongly changing and *spontaneous* creation of photons will occur. Our result indicates that the created photons, when measured far from the star, will carry a net polarization given by ΔQ_D. Some preliminar numerical simulations done by the Valencia numerical relativity group (N. Sanchis-Gual, V. Mewes, J. A. Font) indicate that for a neutron star of $M = 1.73$ solar masses and angular momentum $J = 0.36M^2$ the collapse produces around 30 photons per second more with one circular polarization than the other. This process has no classical counterpart and is different from the standard, late-time Hawking radiation, which does not contribute to ΔQ_D.[14]Although this number is small—given the short duration of the gravitational collapse—it is significant if we compare it with the ≈ 20 total photons per second, steadily emitted by the formed black hole via Hawking radiation, with net polarization equal zero [145].

It is expected that ΔQ_D becomes significantly larger in more violent processes, as for instance the collision and merger of two black holes as the ones observed by the LIGO-Virgo collaboration [5]. A toy-model numerical simulation of two $\sim 0.5 M_\odot$ merging Schwarzchild black holes gives an amount of around 1000 photons, for instance. But more importantly, the existence of spontaneous creation of photons implies that the *stimulated* counterpart must exist. Therefore, electromagnetic radiation traveling in spacetimes with a non-zero value of (5.78), such as the ones mentioned above, will experience a change in its net polarization. Light rays coming from different sides of those systems not only would bend around, but an effective difference in polarization could also be induced between them through quantum fluctuations.

Let us give some additional insights. We need to consider gravitational systems that present both huge tidal (E) and frame dragging (B) effects. As argued astrophysical rotating systems such as stars collapsing to Kerr black holes, or binary mergers in astrophysics, are ideal to address this question. This requires the study of asymptotically

[14] In fact, for an exact Kerr geometry expression (5.78) yields zero if integrated in the whole two-sphere. Therefore, ΔQ_D comes from the (transient) process of collapse, in contrast to the Hawking effect, which is associated with the final, stationary black hole configuration.

flat spacetimes, for which a conformal completion a la Penrose can be made. The use of techniques for dealing with asymptotically flat spacetimes was developed successfully in the 60's by Bondi, Penrose and others [49, 161, 172] in order to determine rigorously the question of gravitational wave emission. See also [30, 99] for comprehensive reviews and more technical and modern details. The physical spacetime M is compactified by adding a boundary $\mathscr{I}^{\pm}$ representing future and past null infinity. Incoming radiation or massless fields start from $\mathscr{I}^{-}$ while outgoing radiation propagate to $\mathscr{I}^{+}$. See Penrose diagram adjoint.

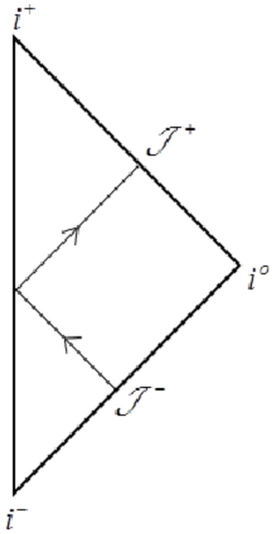

Figure 5.1: Penrose diagram for Minkowski spacetime. Radiation (waves) described by massless fields propagate from past null infinity $\mathscr{I}^{-}$, cross the origin $r = 0$ (vertical line), and propagate then to future null infinity $\mathscr{I}^{+}$.

Given a physical conserved current $\nabla_a j^a_D = 0$, one can deduce that

$$0 = \int_{M \cup \mathscr{I}^+ \cup \mathscr{I}^-} dVol(x) \nabla_a j^a_D(x) = \int_{\mathscr{I}^+} d\mathscr{I}^+_a j^a_D - \int_{\mathscr{I}^-} d\mathscr{I}^-_a j^a_D = Q_+ - Q_- \tag{5.79}$$

where in the second equality we used Stokes Theorem and assumed that no other boundaries exist. The flux of the physical quantity of interest that "enters" spacetime, Q_-, equals the flux of the same physical quantity that leaves, Q_+. This is the standard conservation law. In our case, the Noether charge associated to the classical electromagnetic duality symmetry measures the net difference between right- and left- circularly polarized photons. So this result means, at least classically, that if no electric charges or currents are present in the spacetime, the polarization measured initially at $\mathscr{I}^{-}$ equals that measured finally at $\mathscr{J}^+$. In the astrophysical systems that we are interested in, the appearance of horizons or black hole singularities may require the necessity of taking into account the contribution of the flux through an internal (timelike or null) boundary $\mathscr{I}$ as well:

$$0 = \int_{M\cup\mathscr{J}^+\cup\mathscr{J}^-} dVol(x)\nabla_a j^a_D(x) = Q_+ - Q_- + \int_{\mathscr{I}} d\mathscr{I}_a j^a_D \tag{5.80}$$

On statistical grounds, one could expect the number of right and left photons crossing the horizon to be the same, so the above conclusion could be achieved as well.

In the quantum theory, however, we would have

$$\langle Q_+\rangle - \langle Q_-\rangle + \int_{\mathscr{I}} d\mathscr{I}_a \langle j^a_D\rangle = \int_{M\cup\mathscr{J}^+\cup\mathscr{J}^-} dVol(x)\langle\nabla_a j^a_D(x)\rangle \propto \int_{M\cup\mathscr{J}^+\cup\mathscr{J}^-} dVol(x) E_{ab}B^{ab}$$

and so gravitation can have an influence on the polarization state of photons. Analitically it is difficult (possibly not available) to find solutions to Einstein's equations that give a non-vanishing value of the RHS. So we are forced to solve them numerically in order to give an estimate to this integral. Regarding the LHS, we have to focus on each of the terms individually. In this work we are concerned only with outgoing radiation, i.e. radiation that propagates to future null infinity $\mathscr{J}^+$ and is generated through quantum fluctuations inside the physical spacetime M (i.e. there is no incoming radiation from $\mathscr{J}^-$). Therefore, initially (i.e. at $\mathscr{J}^-$) there are no photons and we can safely say that $\langle Q_-\rangle = 0$ [15]. The calculation of $\langle Q_+\rangle$, however, amounts to obtain

$$\int_{\mathscr{J}^+} d\mathscr{J}^+_a J^a_{CS} \tag{5.81}$$

where J^a_{CS} is the Chern-Simons gravitational current (one can check that $\nabla_a J^a_{CS} = R_{abcd}{}^\star R^{abcd}$). This can be worked out using the above mentioned asymptotic techniques. A calculation following [31] shows that the result depends on the shear σ^0 of the gravitational waves arriving there

$$\int_{\mathscr{J}^+} d\mathscr{J}^+_a J^a_{CS} \sim \int_{\mathscr{J}^+} d\mathscr{J}^+ \mathrm{Im}\left[\ddot{\sigma}^0\dot{\bar{\sigma}}^0\right] \tag{5.82}$$

where dot denotes derivative with respect to retarded time ($u = t - r$). This is to be compared with

$$Q_D \sim \int_{\mathscr{J}^+} d\mathscr{J}^+ \mathrm{Im}\left[\bar{A}^0_2 \dot{A}^0_2\right] \tag{5.83}$$

[15] $\langle Q_-\rangle \neq 0$ if the initial quantum state is not vacuum, but this requires another derivation dealing with the stimulated quantum process, which will be considered in the future.

where A_2^0 is the transverse component of the electromagnetic potential in the Newman-Penrose basis at $\mathscr{I}^+$. This expression is formally equal to the above by identifying the Bondi news $N \sim \dot{\sigma}^0$ with the electromagnetic potential. So perhaps we could interpret the above formula as being the V-stokes parameter associated to gravitational waves. The Bondi news is important because it is known that GW are emitted if and only if $N = 0$. As a consequence, we also conclude that the duality anomaly manifests if gravitational waves are emitted (assuming the internal boundary to be negligible in this discussion).

Finally, the calculation of the flux crossing the horizon is still subject to study, but perhaps it could be related to the transient process of black hole collapse, and could give information about multipole moment expansions of the geometry of the horizon [32]. If so, this anomaly could serve to infer information of the geometry of the black hole horizon by measuring both gravitational and electromagnetic polarization at future null infinity.

We plan to address all these issues in detail in the future.

5.7 Appendices

Appendix A. Noether current

We derive here in more detail the variation of the lagrangian density (5.4) under the infinitesimal transformation (5.3). Namely,

$$\delta\mathscr{L} \;=\; \frac{\partial\mathscr{L}}{\partial A_\nu}\delta A_\nu + \frac{\partial\mathscr{L}}{\partial\nabla_\mu A_\nu}\delta\nabla_\mu A_\nu = -F^{\mu\nu}\nabla_\mu\delta A_\nu = -F^{\mu\nu}\nabla_\mu Z_\nu \tag{5.84}$$

The equality ${}^\star F = dZ + G$ leads to $F = -{}^\star dZ - {}^\star G$. Then ${}^\star G_{\mu\nu}G^{\mu\nu} = ({}^\star F_{\mu\nu} - dZ_{\mu\nu})(-({}^\star dZ)^{\mu\nu} - F^{\mu\nu}) = dZ^{\mu\nu}({}^\star dZ)_{\mu\nu} - F^{\mu\nu\star}F_{\mu\nu} + 2dZ^{\mu\nu}F_{\mu\nu}$, from which we can deduce

$$\delta\mathscr{L} = -\frac{1}{2}\nabla_\mu(A_\nu{}^\star F^{\mu\nu} - Z_\nu{}^\star dZ^{\mu\nu}) - \frac{1}{4}{}^\star G_{\mu\nu}G^{\mu\nu} \tag{5.85}$$

The final term seems to prevents this to be a total derivative. It is, however, vanishing. Given a foliation of spacetime in spatial hypersurfaces with unit timelike normal vector n^μ, the electric field measured by the associated observer is $E_\mu = n^\nu F_{\nu\mu}$. The magnetic part of G is zero, i.e. $n^{\nu\star}\mathrm{G}_{\mu\nu} = 0$, because we work in the reduced phase space in which the source-free Gauss law, $D_\mu E^\mu = 0$, holds [D denoting the 3D Levi-Civita connection]. In fact, Z is introduced as the "potential" for the electric field that solves the constraint, and thus defined by $E^\mu = -\epsilon^{\mu\nu\rho}D_\nu Z_\rho$. Then, since S is antisymmetric, it can be irreducibly decomposed in terms of its electric E_s and magnetic B_s parts as well. The final term on the RHS above is equal to $E_{s,\mu}B_s^\mu$, so this term vanishes. Then the variation of the lagrangian density is a total derivative, $\delta\mathscr{L} = \nabla_\mu h^\mu \sqrt{-g}$.

The Noether current is finally given by

$$j_D^\mu = \frac{1}{\sqrt{-g}}\frac{\partial\mathscr{L}}{\partial\nabla_\mu A_\nu}\delta A_\nu - h^\mu = \frac{1}{2}\left[A_\nu {}^\star F^{\mu\nu} - 2F^{\mu\nu}Z_\nu - {}^\star dZ^{\mu\nu}Z_\nu\right] \tag{5.86}$$

which agrees with (5.5).

Appendix B. The soldering form α^I_{ab}

We dwell here on the formal introduction of the α objects and its algebraic properties, introduced in point 8 of section 5.3, and used throughout the main text.

Remark: the whole derivation can be similarly reproduced for the complex conjugate vector space $\bar{V}$, which is isomorphic to the subspace V_- of anti-self dual 2-forms, $\bar{V} \cong V_-$. All the results are the same, it is only necessary to take the complex conjugate, and denoting with a dot the indices.

1. Definition and properties

Let $\{n^a, x^a, y^a, z^a\}$ be an orthonormal basis in 4D Minkowski spacetime associated with a global family of inertial observers; the metric reads $\eta_{ab} = n_a n_b - x_a x_b - y_a y_b - z_a z_b$. The following set of tensors

$$\alpha^1_{ab} = -2n_{[a}x_{b]} + i2y_{[a}z_{b]}, \tag{5.87}$$

$$\alpha^2_{ab} = -2n_{[a}y_{b]} + i2z_{[a}x_{b]}, \tag{5.88}$$

$$\alpha^3_{ab} = -2n_{[a}z_{b]} + i2x_{[a}y_{b]}, \tag{5.89}$$

constitute an orthogonal basis in the subspace V_+ of complex self-dual 2-forms,

$$\alpha^I_{ab}\alpha^{ab}_J = -4\delta^I_J, \tag{5.90}$$

$$\alpha^I_{ab}\bar{\alpha}^{ab}_J = 0. \tag{5.91}$$

Then, one can write any self-dual two form in this basis, like the electromagnetic field ${}^+F_{ab}$, in terms of 3 components ${}^+H_I$

$${}^+F_{ab} = {}^+H_I \alpha^I_{ab} \tag{5.92}$$

Let V be now a 3-dimensional complex vector space endowed with an euclidean metric η_{IJ} $[I,J=1,2,3]$, and let $\{X^I, Y^I, Z^I\}$ be an orthonormal basis of (V,η). The metric can be expressed as $\eta_{IJ} = -X_I X_J - Y_I Y_J - Z_I Z_J$. We define then a linear isomorphism, "soldering 2-form", between V and V_+, by identifying the corresponding two basis

$$\alpha_{ab}^{\ \ I} = \alpha^1_{ab} X^I + \alpha^2_{ab} Y^I + \alpha^3_{ab} Z^I \tag{5.93}$$

(tensorial products omitted for simplicity) which provides a one-to-one identification between elements of the 3-dimensional complex space V and the subspace of self-dual 2-forms in Minkowski V_+, i.e., $\alpha^i_{ab}\,{}^+H_I = {}^+F_{ab}$, and $\alpha^{ab}_I\,{}^+F_{ab} = 4\,{}^+H_I$. Also note that $i^\star \alpha^{ab}_I = \alpha^{ab}_I$.

For practical purposes it will be convenient to be aware that in the basis above the matrix components of these tensors read:

$$[\alpha]_{ab}^{\ \ I} = -2\left[{}^+\Sigma^{0I}\right]_{ab} \tag{5.94}$$

We can work out the algebraic properties of these tensors in the given basis by using these matrices, and the results will hold for any arbitrary orthonormal basis. For the sake of clarity, we distinguish between the different indices in order to be able to know which basis element corresponds to, although in matrix notation they are simple labels. (note that, as a matrix, the position of the indexes does matter now, even denoted differently)

Proposition. The following properties hold:

$$\begin{aligned}
\alpha_{ab}^{\ \ I}\,\alpha_{cdI} &= -2\left[{}^+\Sigma_{ab}\right]_{cd} && (5.95)\\
\alpha_{abI}\,\alpha^{ab}_{\ \ \ J} &= 4\eta_{IJ} && (5.96)\\
\alpha_{abI}\,\bar{\alpha}^{ab}_{\ \ \ J} &= 0 && (5.97)\\
\alpha^{ab}_{\ \ \ I}\alpha^{c}_{\ \ bJ} &= \eta_{IJ}\eta^{ac} - 2\left[{}^+\Sigma_{IJ}\right]^{ac} && (5.98)
\end{aligned}$$

Proof:

$$\alpha_{ab}^{\ \ I}\,\alpha_{cdI} = \alpha_{ab}^{\ \ I}\alpha_{cd}^{\ \ J}\eta_{IJ} = 4\left[{}^+\Sigma_{0I}\right]_{ab}\left[{}^+\Sigma^{0I}\right]_{cd} = -2\left[{}^+\Sigma_{ab}\right]_{cd} \tag{5.99}$$

$$\begin{aligned}
\alpha_{abI}\,\alpha^{ab}_{\ \ \ J} &= 4\left[{}^+\Sigma_{0I}\right]_{ab}\left[{}^+\Sigma_{0J}\right]^{ab} = 2([\Sigma_{0I}]_{ab} + i\left[{}^\star\Sigma_{0I}\right]_{ab})\left[{}^+\Sigma_{0J}\right]^{ab}\\
&= -4\left[{}^+\Sigma_{0J}\right]_{0I} - 4\left[i^{\star +}\Sigma_{0J}\right]_{0I} = -8\left[{}^+\Sigma_{0J}\right]_{0I} = -4[\Sigma_{0J}]_{0I} = 4\eta_{IJ}
\end{aligned} \tag{5.100}$$

$$\begin{aligned}
\alpha_{abI}\,\bar{\alpha}^{ab}_{\ \ \ J} &= 4\left[{}^+\Sigma_{0I}\right]_{ab}[{}^-\Sigma_{0J}]^{ab} = 2([\Sigma_{0I}]_{ab} + i\left[{}^\star\Sigma_{0I}\right]_{ab})[{}^-\Sigma_{0J}]^{ab}\\
&= -4[{}^-\Sigma_{0J}]_{0I} - 4\left[i^{\star -}\Sigma_{0J}\right]_{0I} = 0
\end{aligned} \tag{5.101}$$

$$\begin{aligned}\alpha^{ab}{}_I\alpha^c{}_{bJ} &= 4\left[{}^+\Sigma_{0I}\right]^{ab}\left[{}^+\Sigma_{0J}\right]^c{}_b = \eta_I^a\eta_J^c + t^a\eta_{IJ}t^c + i\epsilon_{0J}{}^c{}_I t^a + i\epsilon_{0I}{}^a{}_J t^c - \epsilon_{0I}{}^{ab}\epsilon_{0J}{}^c{}_b \\ &= 2i\epsilon_{0J}{}^{[c}{}_I t^{a]} + \eta_{IJ}\eta^{ac} + 2\eta_I^{[a}\eta_J^{c]} = \eta_{IJ}\eta^{ac} - 2\left[{}^+\Sigma_{IJ}\right]^{ac}. \end{aligned} \tag{5.102}$$

The first two expressions reflect the mapping-preservation of the two natural structures of V_+ and V. Namely, $\left[{}^+\Sigma_{ab}\right]_{cd}$ and η_{IJ}. On the other hand, taking the symmetric and antisymmetric parts of the fourth property yields useful equations:

$$\alpha^{[a}{}_{cI}\alpha^{b]c}{}_J = -2\left[{}^+\Sigma^{ab}\right]_{IJ} \tag{5.103}$$

$$\alpha^{(a}{}_{cI}\alpha^{b)c}{}_J = \eta^{ab}\eta_{IJ} \tag{5.104}$$

The quantities $2\left[{}^+\Sigma^{ac}\right]_{IJ}$ are infinitesimal generators of the (1,0) representation, as one can check by direct computation,

$$[2^+\Sigma^{ab}, 2^+\Sigma^{cd}] = \eta^{ac}2[^+\Sigma]^{bd} - \eta^{ad}2[^+\Sigma]^{bc} + \eta^{bd}2[^+\Sigma]^{ac} - \eta^{bc}2[^+\Sigma]^{ad}\ . \tag{5.105}$$

2. Covariant derivative operator

Let us consider now a globally hyperbolic physical spacetime $(M, g_{\mu\nu})$. At each point in this space, we attach a Minkowski vector space through the vierbein $e_a^\mu(x)$. This is a non-coordinate basis that is defined through the isomorphism $g_{\mu\nu}(x) = \eta_{ab}e_\mu^a(x)e_\nu^b(x)$, with a, b some labels and $\eta_{ab} = diag\{+1, -1, -1, -1\}$. At each spacetime point, we regard then $e^a(x)$ as vectors belonging to a certain vector space, that, endowed with η_{ab}, is isomorphic to Minkowski. Latin indices are lowered and raised with η_{ab} while greek indices with $g_{\mu\nu}(x)$. We extend the domain of applicability of the covariant derivative operator to objects belonging to this internal space by imposing the compatibility condition $\nabla_\mu e_\nu^a(x) = 0$. There should be then a connection 1-form ω_μ that satisfies:

$$0 = \nabla_\mu e_a^\nu = \partial_\mu e_a^\nu + \Gamma^\nu_{\mu\alpha}e_a^\alpha + \omega_{\mu a}{}^b e_b^\nu \tag{5.106}$$

The metric-connection compatibility $\nabla_\mu g_{\alpha\beta}(x) = 0$ implies antisymmetry of the connection 1-form, $\omega_\mu^{ab} = \omega_\mu^{[ab]}$. From the previous equation one gets

$$\omega_\mu^{ab} = e_\nu^a\partial_\mu e^{b\,\nu} + \Gamma^\nu_{\mu\alpha}e_\nu^a e^{\alpha\,b} \tag{5.107}$$

In turn, we will also have the space V attached to each point of M. We need to define the extension of the covariant derivative operator when acting on objects of this space. Following [184], we know that the difference between two derivative operators is linear, so that

$$(\nabla_\mu - \bar{\nabla}_\mu)H_I = -C_{\mu I}{}^J H_J\,, \qquad H_I \in V \tag{5.108}$$

One of those can be taken to be the partial derivative operator, $\bar{\nabla}_\mu = \partial_\mu$. We choose the other derivative operator to be compatible with the soldering form, i.e.

$$0 \equiv \nabla_\mu \alpha_I^{\alpha\beta} = \partial_\mu \alpha_I^{\alpha\beta} + \Gamma^\alpha_{\mu\rho}\alpha_I^{\rho\beta} + \Gamma^\beta_{\mu\rho}\alpha_I^{\alpha\rho} - C_{\mu I}{}^J \alpha_J^{\alpha\beta}$$

where we denote $\alpha^{\alpha\beta}{}_I(x) = e^\alpha_a(x) e^\beta_b(x) \alpha^{ab}{}_I$. Taking (5.96) into account we now multiply with $\alpha^I_{\alpha\beta}$,

$$\begin{aligned} C_{\mu I}{}^J &= \frac{1}{2} e^a_\alpha (\partial_\mu e^\alpha_c) \alpha^J_{ab} \alpha^{cb}_I + \frac{1}{2} \alpha^J_{\alpha\beta} \Gamma^\alpha_{\mu\rho} \alpha^{\rho\beta}_I = \frac{1}{2} \alpha^J_{ab} \left[e^a_\alpha (\partial_\mu e^\alpha_c) + \Gamma^\nu_{\mu\alpha} e^a_\nu e^\alpha_c \right] \alpha^{cb}_I \\ &= \frac{1}{2} \alpha^J_{ab} \alpha_{Ic}{}^b \omega^{ac}_\mu = \frac{1}{2} \alpha^{[a}{}_b{}^J \alpha^{c]b}{}_I \,\omega_{\mu ac} = \frac{1}{2} \omega_\mu{}^{ab} 2 \left[{}^+\Sigma_{ab} \right]_I{}^J \end{aligned} \tag{5.109}$$

Therefore, the covariant derivative acting on the field H^+_I is given by

$$\nabla_\mu H^+_I = \partial_\mu H^+_I - \frac{1}{2} \omega^{ab}_\mu 2 [{}^+\Sigma_{ab}]_I{}^J H^+_J \,. \tag{5.110}$$

We note that the general form of the covariant derivative is given by $\nabla_\mu = \partial_\mu - \frac{1}{2}\omega^{ab}_\mu \Sigma_{ab}$, where Σ_{ab} stands for the generator of the representation associated to the field to be derived. For instance, for the field H^+_I, the generator Σ_{ab} is $2[{}^+\Sigma_{ab}]_I{}^J$ and hence the covariant derivative is (5.110).

One can check immediately the following identities involving the covariant derivative of invariant tensors

$$\begin{aligned} \nabla_\mu \eta_{IJ} &= -\omega_\mu{}^{ab} \left[{}^+\Sigma_{ab} \right]_I{}^K \eta_{KJ} - \omega_\mu{}^{ab} \left[{}^+\Sigma_{ab} \right]_J{}^K \eta_{IK} \\ &= -2\omega_\mu{}^{ab} \left[{}^+\Sigma_{ab} \right]_{(IJ)} = 0 \end{aligned} \tag{5.111}$$

and the same for η^{IJ} or δ^I_J. Secondly $[\epsilon^{IJK}\epsilon_{IJK} = -6]$

$$\begin{aligned} 0 &= \nabla_\mu \epsilon_{IJK} \\ &= -\omega_\mu{}^{ab} \left[{}^+\Sigma_{ab} \right]_I{}^L \epsilon_{LJK} - \omega_\mu{}^{ab} \left[{}^+\Sigma_{ab} \right]_J{}^L \epsilon_{ILK} - \omega_\mu{}^{ab} \left[{}^+\Sigma_{ab} \right]_K{}^L \epsilon_{IJL} \end{aligned}$$

which can be checked to vanish identically by recalling the identity

$$\eta_{I[N}\epsilon_{M]JK} + \eta_{J[N}\epsilon_{M]KI} + \eta_{K[N}\epsilon_{M]IJ} = 0 \,. \tag{5.112}$$

3. 3+1 spacetime decomposition and the γ^μ_I

If spacetime is globally hyperbolic it can be foliated by a family of spatial hypersurfaces Σ_t, $M \simeq \mathbb{R} \times \Sigma_t$, all of them orthogonal to a unit time-like vector n^μ at each point (see [184]). The metric can be decomposed as $g_{\mu\nu} = n_\mu n_\nu + h_{\mu\nu}$, where $h_{\mu\nu}$ is the induced spatial metric (the projected metric tensor) on the hypersurfaces.

We can now use the isomorphism (5.93) to build the following mixed tensors:

- $\gamma_I^{\mu} := n_\nu \alpha^{\nu\mu}{}_I$ provides an isomorphisms between vectors in V and vectors in space-time that are spatial (in the inertial frame we are using).

- $\gamma_{\dot{I}}^{\mu} := n_\nu \bar{\alpha}^{\nu\mu}{}_{\dot{I}}$ is similar to the previous map replacing V by its complex conjugated space $\bar{V}$

- $\gamma_I^{\dot{I}} := \gamma_I^{\mu} \gamma_\mu^{\dot{I}}$ provides an isomorphisms between V and $\bar{V}$.

- $\epsilon^{I\mu\nu} := n_\alpha \gamma_\beta^I \epsilon^{\alpha\beta\mu\nu}$ defines a totally antisymmetric, purely spatial tensor with mixed indces.

In particular, from the last definition one can derive (use $\epsilon_{\alpha\beta\mu\nu} = i\left[{}^+\Sigma_{\alpha\beta}\right]_{\mu\nu} - i\left[{}^-\Sigma_{\alpha\beta}\right]_{\mu\nu}$)

$$i\,2\epsilon^{I\mu\nu} = \alpha^{\mu\nu I} - \bar{\alpha}^{\mu\nu\dot{J}}\,\delta_{\dot{J}}^{I} \tag{5.113}$$

As we have just seen, there is a one-to-one correspondence between ${}^+F_{\mu\nu}$ and ${}^+H_I$. Now we can see that the associated 1-form ${}^+H_\mu = n^{\nu+}F_{\nu\mu}$ is also in one-to-one correspondence. Indeed:

$$H_\nu^+ \;=\; n^{\mu+}F_{\mu\nu} = n^\mu \alpha_{\mu\nu}^{\;\;I} H_I^+ = \gamma_\nu^I H_I^+ \tag{5.114}$$

where we introduced $\gamma_\nu^I := \alpha_{\mu\nu}^I n^\mu$. From the above definitions, and using (5.95) and (5.98), one can easily verify the following properties,

$$\begin{aligned}\gamma_I^\nu \gamma^{\beta I} &= \alpha_I^{\mu\nu} n_\mu \alpha^{\rho\beta I} n_\rho = -2\left[{}^+\Sigma^{\mu\nu}\right]^{\rho\beta} n_\mu n_\rho = -\left[\Sigma^{\mu\nu}\right]^{\rho\beta} n_\mu n_\rho \\ &= -n^\nu n^\beta + g^{\nu\beta} = h^{\nu\beta} \qquad (5.115)\\ \gamma_I^\nu \gamma_{\nu J} &= \alpha_I^{\mu\nu} n_\mu \alpha^{\rho\nu J} n_\rho = \alpha_I^{(\mu|\nu} n_\mu \alpha^{\rho)\nu J} n_\rho \\ &= \eta_{IJ} \qquad (5.116)\end{aligned}$$

Notice that $\nabla_\mu \gamma_I^\nu \neq 0$, but $D_\mu H_+^\mu = D_I H_+^I$ holds true. Proof:

$$\begin{aligned}D_\mu H_+^\mu &= h^{\mu\nu}\nabla_\mu H_\nu^+ = h^{\mu\nu}\nabla_\mu \gamma_\nu^I H_I^+ = \gamma^{\mu I}\nabla_\mu H_I^+ + h^{\mu\nu}(\nabla_\mu \gamma_\nu^I) H_I^+ \\ &= \gamma^{\mu I}\nabla_\mu H_I^+ + h^{\mu\nu}\alpha^{\sigma}{}_{\nu}{}^{I}(n_\mu a_\sigma + K_{\mu\sigma}) H_I^+ = \gamma^{\mu i}\nabla_\mu H_I^+ \equiv D_I H_+^I \qquad (5.117)\end{aligned}$$

where we used that $\nabla_\mu n_\sigma = n_\mu a_\sigma + K_{\mu\sigma}$, a_σ is the 4-acceleration of n_ν and $K_{\mu\nu} = K_{(\mu\nu)}$ its extrinsic curvature.

4. α^I_{ab} as an invariant symbol of the Lorentz group

Consider a local Lorentz transformation Λ acting on the self-dual field $F^+_{ab} = H^+_I \alpha_{ab}^{\ I}$ with the tensorial rule

$$F^+ \to F'^+ = \Lambda^T F^+ \Lambda = H'^+_I \alpha^I \ . \tag{5.118}$$

By construction this transformation leave invariant the quantity

$$-\frac{1}{4} F'^{+ab} F'^+_{ab} = -\frac{1}{4} F^{+ab} F^+_{ab} \ , \tag{5.119}$$

which corresponds to $H'^+_I H'^I_+ = H^+_I H^I_+$. In other words, the Lorentz transformation Λ induces an action

$$H^+_I \to H'^+_I = R(\Lambda)_I^{\ J} H^+_J \tag{5.120}$$

preserving the metric η_{IJ} of the three-dimensional complex space. This is just the $(1,0)$ representation of the Lorentz group. The above matrices $\alpha_{ab}^{\ I}$ define then an invariant symbol, and hence an invariant mixed tensor, of the Lorentz group:

$$R^I_J \alpha_{cd}^{\ \ J} \Lambda^c_{\ a} \Lambda^d_{\ b} = \alpha_{ab}^{\ \ I} \ . \tag{5.121}$$

The symbol η_{IJ} is also an invariant symbol of the Lorentz group:

$$\eta_{MN} R^M_{\ I} R^N_{\ J} = \eta_{IJ} \ . \tag{5.122}$$

It can be regarded as the analogue of the metric η_{ab} in the three-dimensional complex space of the $(1,0)$ representation. Note that it is also the analogue of the invariant (antisymmetric) symbol $\epsilon_{\alpha\beta}$ of the $(1/2,0)$ representation. Moreover, since the matrices $R(\Lambda)_I^{\ J}$ have unit determinant, the Levi-Civita symbol ϵ^{IJK} is also an invariant tensor of the Lorentz group.

In the spinor language of [33, 100], the basic tensor $\alpha_{ab}^{\ I}$ can also be regarded as a "soldering form", i.e., an isomorphism identifying the space of self-dual two-forms in $4D$ Minkowski space and the three-dimensional complex space with an Euclidean scalar product.

We can also construct the $(0,1)$ Lorentz representation using the anti-self-dual field strength and the matrices $\bar{\alpha}_{ab}^{\ \dot{J}}$. With the definition

$$F^-_{ab} = H^-_{\dot{J}} \alpha_{ab}^{\ \dot{J}} \ , \tag{5.123}$$

we can define the Lorentz action

$$F^- \to F'^- = \Lambda^T F^- \Lambda = H'^-_{\dot{I}} \bar{\alpha}^{\dot{I}} \ , \tag{5.124}$$

and therefore

$$H^-_{\dot{I}} \to H'^-_{\dot{I}} = R(\Lambda)_{\dot{I}}^{\ \dot{J}} H^-_{\dot{J}} \ . \tag{5.125}$$

Note that $R(\Lambda)_{\dot{I}}^{\ \dot{J}} = (R(\Lambda)_I^{\ J})^*$. The invariant symbols for the $(0,1)$ representation are

$$R^{\dot{I}}_{\ \dot{J}} \bar{\alpha}_{cd}^{\ \ \dot{J}} \Lambda^c_{\ a} \Lambda^d_{\ b} = \bar{\alpha}_{ab}^{\ \ \dot{I}} \ , \tag{5.126}$$

$$\eta_{\dot{M}\dot{N}} R^{\dot{M}}_{\ \dot{I}} R^{\dot{N}}_{\ \dot{J}} = \eta_{\dot{I}\dot{J}} \ , \tag{5.127}$$

where

$$\epsilon_{\dot{M}\dot{N}\dot{O}} R^{\dot{M}}_{\ \dot{I}} R^{\dot{N}}_{\ \dot{J}} R^{\dot{O}}_{\ \dot{K}} = \epsilon_{\dot{I}\dot{J}\dot{K}} \ . \tag{5.128}$$

5. Extended α^I_{ab} and its algebraic properties

In order to fix the gauge in the first-order Lagrangian (5.47) in section 5.4 we extended the internal vector space structure to accommodate for the Lagrange multiplier $H^0_{\pm}$. We introduced then a fourth α element, as normally done with the spin$-1/2$ case with Pauli matrices. We show here the new α elements and its extended algebraic properties, that are used throughout the main text.

Let $\hat{V} = V \oplus \mathbb{R}$, endowed with a Lorentzian flat metric η_{IJ}, be now our internal vector space at each spacetime point, with indices I, J running from 0 to 3. The complex 3-dimensional vector space V, isomorphic to the space of self-dual 2-forms in Minkowski, is now a vector subspace. Let n_I denote a unit vector ($\eta_{IJ} n^I n^J = 1$) orthogonal to the V subspace ($n^I m^I \eta_{IJ} = 0$, $m^J \in V$). It spans a 1-dimensional vector space. In fact the metric in this space can be written as $\eta_{IJ} = n_I n_J + \gamma_{IJ}$. By defining $n^I \alpha^{\mu\nu}_I := -g^{\mu\nu}$, we can write

$$\alpha^{\mu\nu}_I = -n_I g^{\mu\nu} + \alpha^{\mu\nu}_J \gamma^J_I \tag{5.129}$$

where we denote now $\gamma^J_I := \gamma^\nu_I \gamma_{\nu J}$. Notice that $\alpha^{\mu\nu}_J \gamma^J_I$ is now what in previous subsections of this Appendix we identified as $\alpha^{\mu\nu}_I$ with $I = 1,2,3$.

As argued before, in order to define the actuation of the covariant derivative operator of elements belonging to this (extended) internal vector space we demand the compatibility condition

$$\nabla_\rho \alpha^{\mu\nu}_I = 0 \tag{5.130}$$

and equivalently $\nabla_\rho \bar{\alpha}^{\mu\nu}_{\dot{I}} = 0$. We notice that this condition, in turn, imply several other useful ones. First of all, by following a similar reason as (5.111) we conclude that

$\nabla_\mu \eta_{IJ} = 0$. Secondly, by multiplying with the metric $g_{\mu\nu}$ the above expression using (5.129) one gets $\nabla_\rho n_I = 0$. But turning back again to (5.130) one finds the old compatibility condition:

$$\nabla_\rho(\alpha^{\mu\nu}_{J}\gamma^{J}_{I}) = 0 \tag{5.131}$$

Therefore, all results that used this condition are still valid with the new extended internal space and new compatibility condition. Finally, if $\nabla_\mu n_I = 0$ and $\nabla_\mu \eta_{IJ} = 0$, then $\nabla_\mu \gamma_{IJ} = 0$.

Now we shall derive the algebraic properties of the new alpha matrices when the index I runs from 0 to 3. Equation (5.95) can be written also as $\alpha^{\ I}_{ab}\alpha^{\ J}_{cd}\gamma_{IJ} = -2[^{-}\Sigma_{ac}]_{bd} - [\Sigma_{ad}]_{cb}$, and then

$$\alpha^{\ I}_{ab}\alpha_{cdI} = \alpha^{\ I}_{ab}\alpha^{\ J}_{cd}\gamma_{IJ} + g_{ab}g_{dc} = \eta_{bd}\eta_{ac} - 2[^{-}\Sigma_{ac}]_{bd} \tag{5.132}$$

Furthermore, from (5.96) we can obtain

$$\alpha_{abI}\alpha^{ab}_{\ \ J} = \alpha_{abM}\alpha^{ab}_{\ \ N}\gamma^{M}_{I}\gamma^{N}_{J} + \eta_{ab}\eta^{ab}n_I n_J = 4\gamma_{IJ} + 4n_I n_J = 4\eta_{IJ} \tag{5.133}$$

The third property, (5.98), yields

$$\begin{aligned}
\alpha^{ab}_{\ \ I}\alpha^{c}_{\ bJ} &= \alpha^{ab}_{\ \ M}\alpha^{c}_{\ bN}\gamma^{M}_{I}\gamma^{N}_{J} - n_I\alpha^{ca}_{\ \ J} - n_J\alpha^{ac}_{\ \ I} + \eta^{ac}n_I n_J \\
&= \alpha^{ab}_{\ \ M}\alpha^{c}_{\ bN}\gamma^{M}_{I}\gamma^{N}_{J} + \eta^{ac}n_I n_J + 2n_I[^{+}\Sigma_{0J}]^{ca} + 2n_J[^{+}\Sigma_{I0}]^{ca} \\
&= \eta_{IJ}\eta^{ac} - 2\left[^{+}\Sigma^{ac}\right]_{IJ}
\end{aligned} \tag{5.134}$$

Finally, we shall prove that, as matrices, the alphas are hermitian, $(\alpha^\mu)^\dagger = \alpha^\mu$.

$$\begin{aligned}
\left[\alpha^{a}\right]^{b}_{\ I} &= -n_I g^{ab} - 2[^{+}\Sigma_{0I}]^{ab} = -n_I g^{ab} - 2[^{-}\Sigma^{0a}]^{\ b}_{I} - [\Sigma^{0b}]^{\ a}_{I} \\
&= -n_I n^a n^b - n_I h^{ab} + i\epsilon^{ab}_{\ \ I} - h^{a}_{I}n^b + h^{ab}n_I - h^{ab}n_I + h^{b}_{I}n^a \\
&= -n_I n^a n^b - n_I h^{ab} - h^{a}_{I}n^b - h^{b}_{I}n^a + i\epsilon^{ab}_{\ \ I}
\end{aligned} \tag{5.135}$$

with this expression it is clear that $[\alpha^a]^{b}_{\ I} = [\bar{\alpha}^a]^{\ b}_{I}$.

Appendix C. Maxwell equations in curved spacetime

In this appendix we shall be concerned on deriving the first-order equations of motion for the vector potentials in curved space, which was presented in Sec. IIIB. Although conceptually equal to the Minkowskian case, Sec. IIIA, it is technically more involved, and we give the details here.

The full set of Maxwell equations $\nabla_\mu F^{\mu\nu}_{\pm} = 0$ can be rewritten in the Weyl form

$$\alpha^{\mu\nu}_{\ \ I} \nabla_\mu H^I_+ = 0\ , \qquad \bar{\alpha}^{\mu\nu}_{\ \ \dot{I}} \nabla_\mu H^{\dot{I}}_- = 0\ , \tag{5.136}$$

Given a spacetime foliation, take t^μ as the unit timelike vector field normal to the hypersurface Σ_t. We can solve the constraint equations, i.e.,

$$D_I H^I_+ \equiv n_\nu (\alpha^\mu)^\nu_{\ I} \nabla_\mu H^I_+ \quad = \quad 0 \tag{5.137}$$

(and its complex conjugate) to introduce complex potentials by means of the "curl"

$$H^I_\pm = \pm i \epsilon^{I\mu\nu} \nabla_\mu A^\pm_\nu \tag{5.138}$$

where we define $\epsilon^{I\mu\nu}$ and γ^ν_I as in 5.7. Indeed,

$$D_I H^I_+ = D_\alpha H^\alpha_+ = i D_\alpha \epsilon^{\alpha\mu\nu} \nabla_\mu A^+_\nu = \epsilon^{\alpha\mu\nu} D_\alpha \nabla_\mu A^+_\nu = \epsilon^{\alpha\mu\nu} \nabla_\alpha \nabla_\mu A^+_\nu \propto \epsilon^{\alpha\mu\nu} R_{\alpha\mu\nu\beta} A^\beta_+ = 0 \tag{5.139}$$

where in the first equality we used the identity derived in (5.117); also (5.114) and (5.115). Note that the potentials inherit the usual gauge ambiguity $A^+_\nu \to A^+_\nu + \nabla_\mu \lambda$. The second-order equations of motion for these potentials can be obtained form (5.136). The projection with n^μ is trivial by construction, while the remaining components satisfy the standard wave equations

$$h^\alpha_\beta [\Box A^\beta_+ - \nabla_\nu \nabla^\beta A^\nu_+] = 0 \tag{5.140}$$

where we made use of (5.113). There are only 3 independent dynamical equations, while the fourth one is the (Gauss) constraint used to define the complex potentials. Note also that we have not fixed the gauge freedom so far.

We are interested in getting a closer view to Dirac spin 1/2 theory, in particular, working with a first-order Lagrangian. Since these equations of motion are of second order, one may expect that source-free Maxwell action can only be of second-order for potentials. We shall see right now that, however, that these second order equation can be obtained from first-order ones thanks to the introduction of complex potentials.

Maxwell equations in the absence of charges and currents can be recovered from $d\,{}^+F = 0$, and requiring ${}^+F_{ab}$ be self-dual, i.e. , ${}^+F_{ab} = \frac{i}{2} \epsilon_{abcd}\, {}^+F^{cd}$, which in turn implies the irreducible decomposition ${}^+F_{ab} = \frac{1}{2}\left[F_{ab} + i\, {}^\star F_{ab}\right]$ for some 2-form F. Indeed, both expressions lead to

$$dF = 0, \qquad d\, {}^\star F = 0. \tag{5.141}$$

and hence F is an electromagnetic field. Conversely, given these equations for an electromagnetic field F, one can construct ${}^{\pm}F := \frac{1}{2}[F \pm i{}^{\star}F]$, which satisfies ${}^{+}F_{ab} = \frac{i}{2}\epsilon_{abcd}{}^{+}F^{cd}$ and $d\,{}^{+}F = 0$.

We shall take this reformulation as our starting point. We shall always assume ${}^{+}F$ as a self-dual 2-form. The equation $d\,{}^{+}F = 0$ allows one to define the local potential A_+ by writing ${}^{+}F = dA_+$; while the self-duality property ${}^{+}F_{ab} = \frac{i}{2}\epsilon_{abcd}{}^{+}F^{cd}$, when written in terms of A_+, gives the dynamical equations (5.41) for A_+:

$$\nabla_a A_b^+ - \nabla_b A_a^+ = i\epsilon_{abcd}\nabla^c A_+^d \tag{5.142}$$

The first-order equation of motion for the 4-vector potential can be rearranged as

$$[{}^{-}\Sigma^{\alpha\beta}]^{\mu\nu}\nabla_\mu A_\nu^+ = 0\,. \tag{5.143}$$

where $[{}^{-}\Sigma^{\alpha\beta}]^{\mu\nu} = \frac{1}{2}\{[\Sigma^{\alpha\beta}]^{\mu\nu} - i[{}^{\star}\Sigma^{\alpha\beta}]^{\mu\nu}\}$ with $[\Sigma^{\alpha\beta}]^{\mu\nu} = \frac{1}{2}\epsilon^{\alpha\beta\rho\sigma}\epsilon_{\rho\sigma}{}^{\mu\nu}$, $[{}^{\star}\Sigma^{\alpha\beta}]^{\mu\nu} = -\epsilon^{\alpha\beta\mu\nu}$.

Now, we shall show that this equation is equivalent to the standard second-order equation of motion for the 4-vector potential. Acting with the derivative operator

$$[{}^{-}\Sigma^{\alpha\beta}]^{\mu\nu}\nabla_\alpha\nabla_\mu A_\nu^+ = 0 \tag{5.144}$$

leads to

$$\nabla^\nu\nabla^\beta A_\nu^+ - \Box A_+^\beta + i\epsilon^{\alpha\beta\mu\nu}\nabla_\alpha\nabla_\mu A_\nu^+ = 0 \tag{5.145}$$

The second term vanishes because of the Bianchi Identity:

$$\epsilon^{\alpha\beta\mu\nu}\nabla_{[\alpha}\nabla_{\mu]}A_\nu^+ \propto \epsilon^{\alpha\beta\mu\nu}R_{\alpha\mu\nu\rho}A_+^\rho = 0 \tag{5.146}$$

Thus we get

$$\Box A_+^\beta - \nabla_\nu\nabla^\beta A_+^\nu = 0 \tag{5.147}$$

which is precisely what one gets from $\nabla_\mu{}^{+}F^{\mu\nu} = 0$ if ${}^{+}F = dA^+$. The same is true for the antiself dual potential, $\nabla_\mu{}^{-}F^{\mu\nu} = 0$. Since ${}^{+}F_{ab} = \frac{1}{2}\left[F_{ab} + i{}^{\star}F_{ab}\right]$, joining both expressions together we end up with

$$\nabla_\mu F^{\mu\nu} = 0\,, \quad \nabla_\mu{}^{\star}F^{\mu\nu} = 0\,. \tag{5.148}$$

Note that there are 16 equations in (5.143), but only 3 are linearly independent (antisymmetry reduces to 6, and anti-selfduality brings down this number to 3). To better express the dynamics act on the LHS by $[\bar{\alpha}_\alpha]_\beta{}^{\dot{I}}$, and using properties (5.95) and (5.96) we get

$$\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu A^+_\nu = 0\,, \qquad \alpha^{\mu\nu}{}_I\nabla_\mu A^-_\nu = 0\,. \tag{5.149}$$

From this equation we see that A_0 has no dynamics (no time derivative for it appears). There are then 3 (complex) first-order differential equations for 3 (complex) propagating degrees of freedom in phase space. Since there is gauge freedom (only the transversal part of the potentials contribute non-trivially to the equations), one can supplement this purely dynamical equation with a gauge fixing.

We put emphasis in the fact that the the duality condition of the electromagnetic field gives source-free Maxwell equations, and that complex potentials allow the diagonalization of them, allowing to write first-order equations of motion. It is worth to remark that in Minkowski one can find an alternative derivation of this equations using vector calculous identities, as did in the main text.

Alternative derivation: on-shell $F^+ = dA^+$ holds, so inverting (5.92) using (5.96) we can also find useful relations:

$$H^+{}_I \;=\; \frac{1}{4}\alpha^{ab}{}_I\;{}^+F_{ab} = \frac{1}{2}\alpha^{\mu\nu}{}_I\;\nabla_\mu A^+_\nu \tag{5.150}$$

but notice this is only valid at the level of equations of motion. In fact, by using the off-shell valid relations (5.138), together with (5.113), we recover in a different way equations (5.149) for complex potentials.

Appendix D. Maxwell action from the first-order Lagrangian in self-dual variables

We prove here that the action functional (5.47) is equivalent to the standard Maxwell action (5.1). We do the calculation in Minkowski, but the generalization to curved spacetimes is straightforward.

1. First of all we prove here that the symplectic structure with self-dual variables is equivalent to the standard one, i.e., that the transformation $\{A_+, A_-\} \to \{A; -E\}$ is canonical.

$$\begin{aligned} -\int d^3\vec{x}[\delta_1 A^i_+\delta_2 H^-_i - \delta_2 A^i_+\delta_1 H^-_i] \;&\equiv\; -\int d^3\vec{x}\,\delta A^i_+ \wedge \delta H^-_i \to -\frac{1}{2}\int d^3\vec{x}(\delta A + i\delta Z)^i \wedge (\delta E - i\delta B)_i \\ &=\; -\frac{1}{2}\int d^3\vec{x}\{\delta A^i \wedge \delta E_i + \delta Z^i \wedge \delta B_i + i\delta Z^i \wedge \delta E_i - i\delta A^i \wedge \delta B_i\} \end{aligned}$$

Since $\vec{B} = \vec{\nabla} \times \vec{A}$ and $\vec{E} = -\vec{\nabla} \times \vec{Z}$, by integrating by parts the last two terms vanish. Now:

$$\begin{aligned} -\int d^3\vec{x}\,\delta A^i_+ \wedge \delta H^-_i &\to -\frac{1}{2}\int d^3\vec{x}\,\delta A^i \wedge \delta E_i + \delta Z^i \wedge \epsilon_{ijk}\partial^j \delta A^k = -\frac{1}{2}\int d^3\vec{x}\,\delta A^i \wedge \delta E_i - \epsilon_{ijk}\partial^j \delta Z^i \wedge \delta A \\ &= -\int d^3\vec{x}\,\delta A^i \wedge \delta E_i \end{aligned} \quad (5.15$$

Exactly the same analysis can be carried out with the other couple (A_-, H_+). Then

$$-\int d^3\vec{x}\,\delta A^i \wedge \delta E_i \to -\frac{1}{2}\int d^3\vec{x}\,\{\delta A^i_+ \wedge \delta H^-_i + \delta A^i_- \wedge \delta H^+_i\} \quad (5.152)$$

2. We prove now that the first-order action written in terms of self-dual variables (5.47) is equivalent to the standard one. Namely,

$$L[\vec{A}_+, \vec{A}_-, \dot{\vec{A}}_+, \dot{\vec{A}}_-] = -\frac{1}{2}\int d^3\vec{x}\{\dot{\vec{A}}^i_+ H^-_i + \dot{\vec{A}}^i_- H^+_i\} - H[A_+, A_-] \quad (5.153)$$

with Hamiltonian

$$H[A_+, A_-] = \int d^3\vec{x}(\vec{\nabla} \times \vec{A}_+)(\vec{\nabla} \times \vec{A}_-) + \frac{1}{2}A^+_0 \partial_i H^i_- + \frac{1}{2}A^-_0 \partial_i H^i_+ \quad (5.154)$$

when we make the transformation $\{A_+, A_-\} \to \{A, ; -E\}$ it leads to

$$\begin{aligned} \int dt L[\vec{A}_+, \vec{A}_-, \dot{\vec{A}}_+, \dot{\vec{A}}_-] &= -\int d^4x Re\,\dot{A}^i_+ H^-_i - \int dt H[A_+, A_-] \\ &\to -\frac{1}{2}\int d^4x Re(\dot{A} + i\dot{Z})^i (E - iB)_i - \int dt H[A_+, A_-] \\ &= -\frac{1}{2}\int d^4x \dot{A}^i E_i + \dot{Z}^i B_i - \int dt H[A_+, A_-] \\ &= -\frac{1}{2}\int d^4x \dot{A}^i E_i + \dot{Z}^i \epsilon_{ijk}\partial^j A^k - \int dt H[A_+, A_-] \\ &= -\frac{1}{2}\int d^4x \dot{A}^i E_i - Z^i \epsilon_{ijk}\partial^j \dot{A}^k - \int dt H[A_+, A_-] \\ &= -\frac{1}{2}\int d^4x \dot{A}^i E_i + \epsilon_{ijk}\partial^j Z^i \dot{A}^k - \int dt H[A_+, A_-] \\ &= -\int d^4x \dot{A}^i E_i - \int dt H[A_+, A_-] \end{aligned} \quad (5.155)$$

where in the process we integrated by parts in time and space. Notice that in this transformation

$$H[A_+, A_-] = \int d^3\vec{x} H^i_- H^+_i \to \frac{1}{2}\int d^3\vec{x}(E^2 + (\epsilon_{ijk}\partial^j A^k)^2) = H[A, E] \quad (5.156)$$

so that in second order or lagrangian formalism [after the Legendre transform which takes E_i to $\dot{A}_i$] we recover the standard result

$$\int dt L[\vec{A}_+, \vec{A}_-, \dot{\vec{A}}_+, \dot{\vec{A}}_-] \to \int dt L[A, \dot{A}] = \int d^4x\{\frac{1}{2}(\epsilon^{ijk}\partial_j A_k)^2 - \frac{1}{2}\dot{A}_i \dot{A}^i\} = -\frac{1}{4}\int d^4x F_{ab}F^{ab} \quad (5.157)$$

3. Now we prove that this first-order Lagrangian (5.153) can be written in a manifestly Lorentz invariant way. From (5.47) and (5.43) we can deduce

$$\begin{aligned} L(\vec{A}_+,\vec{A}_-,\dot{A}_+,\dot{A}_-) &= -\frac{1}{2}\int d^3\vec{x}\{H^-_i(\dot{A}^i_+ - i\epsilon^{ijk}\partial_j A^+_k) + H^+_i(\dot{A}^i_- + i\epsilon^{ijk}\partial_j A^-_k) + A^+_0\partial_i H^i_- + A^-_0\partial_i H^i_+\} \\ &= -\frac{1}{2}\int d^3\vec{x}\{H^i_-\bar{\alpha}^{\mu j}{}_i\partial_\mu A^+_j + H^i_+\alpha^{\mu j}{}_i\partial_\mu A^-_j + A^+_0\partial_i H^i_- + A^-_0\partial_i H^i_+\} \\ &= -\frac{1}{2}\int d^3\vec{x}\{H^i_-\bar{\alpha}^{\mu\nu}{}_i\partial_\mu A^+_\nu + H^i_+\alpha^{\mu\nu}{}_i\partial_\mu A^-_\nu\} \\ &\equiv -\frac{1}{2}\int d^3\vec{x}\{H^{\dot{I}}_-\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\partial_\mu A^+_\nu + H^I_+\alpha^{\mu\nu}{}_I\partial_\mu A^-_\nu\} \end{aligned} \tag{5.158}$$

where in the last step we introduced the appropriate notation to indicate the behaviour of the indices under Lorentz transformations (this is the correct way to say that this expression is Lorentz invariant). See Appendix B for details regarding the α matrices. The generalization to curved spacetime is straightforward from here. In particular, one can directly invoke the minimal coupling prescription from (5.158) and gets the second line in (5.47).

Appendix E. Deriving the equations of motion from the first-order action

We check here that Maxwell equations for the complex potentials are derived correctly. Given the action functional (5.47)

$$S_M[A_+,A_-] = -\frac{1}{2}\int d^4x\sqrt{-g}\left[H^{\dot{I}}_-\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu A^+_\nu + H^I_+\alpha^{\mu\nu}{}_I\nabla_\mu A^-_\nu\right] \tag{5.159}$$

we recover Maxwell equations. Indeed,

$$0 = \frac{\delta S_M}{\delta A^+_\nu} = \frac{1}{2}\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu H^{\dot{I}}_- + \nabla_\mu\frac{i}{2}\epsilon^{I\mu\nu}\alpha^{\alpha\beta}{}_I\nabla_\alpha A^-_\beta \tag{5.160}$$

we use now the identity (5.113) [note that since $\nabla_\mu\bar{\alpha}^{\mu\nu I}\neq 0$, then $\nabla_\mu\epsilon^{I\mu\nu}\neq 0$] and (5.95) to write

$$\begin{aligned} 0 &= \frac{1}{2}\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu H^{\dot{I}}_- + \frac{1}{4}\alpha^{\mu\nu I}\alpha^{\alpha\beta}{}_I\nabla_\mu\nabla_\alpha A^-_\beta - \frac{i}{2}\nabla_\mu\bar{\alpha}^{\mu\nu I}\epsilon^{\alpha\beta}{}_I\nabla_\alpha A^-_\beta - \frac{1}{4}\bar{\alpha}^{\mu\nu I}\bar{\alpha}^{\alpha\beta}{}_{\dot{I}}\nabla_\mu\nabla_\alpha A^-_\beta \\ &= \frac{1}{2}\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu H^{\dot{I}}_- - \frac{i}{2}\nabla_\mu\bar{\alpha}^{\mu\nu\dot{I}}\epsilon^{\alpha\beta}{}_{\dot{I}}\nabla_\alpha A^-_\beta - \frac{1}{2}[^+\Sigma^{\mu\nu}]^{\alpha\beta}\nabla_\mu\nabla_\alpha A^-_\beta + \frac{1}{2}[^-\Sigma^{\mu\nu}]^{\alpha\beta}\nabla_\mu\nabla_\alpha A^-_\beta \end{aligned} \tag{5.161}$$

[at certain points we changed I to $\dot{I}$ like in $\bar{\alpha}^{\mu\nu I}\bar{\alpha}^{\alpha\beta}{}_I = \bar{\alpha}^{\mu\nu\dot{I}}\bar{\alpha}^{\alpha\beta}{}_{\dot{I}}$ in order to take the derivative operator ∇_μ through]. Recalling that $H^{\dot{I}}_- = -i\epsilon^{\dot{I}\mu\nu}\nabla_\mu A^\pm_\nu$, and using the Bianchi

Identity $\epsilon^{abcd}R_{bcde} = 0$, we get

$$\begin{aligned} 0 &= \bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu H_-^{\dot{I}} - i[{}^\star\Sigma^{\mu\nu}]^{\alpha\beta}\nabla_\mu\nabla_\alpha A_\beta^- \\ &= \bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu H_-^{\dot{I}} + \frac{i}{2}\epsilon^{\mu\nu\alpha\beta}R_{\mu\alpha\beta\sigma}A_-^\sigma \\ &= \bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu H_-^{\dot{I}} \end{aligned} \tag{5.162}$$

Analogously, by differentiating with respect to A_- we get the corresponding complex-conjugate equation.

This particular calculation allows to write the action functional (5.47) as

$$S_M[A_+, A_-] = -\int d^4x\sqrt{-g}\, H_-^{\dot{I}}\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu A_\nu^+ \tag{5.163}$$

since, by integrating by parts and assuming the fields to decay sufficiently fast at infinity so that we can neglect "boundary" terms, one can express

$$-\frac{1}{2}\int d^4x\sqrt{-g}\, H_+^I\alpha^{\mu\nu}{}_I\nabla_\mu A_\nu^- = \int d^4x\sqrt{-g}\,\frac{i}{2}A_\nu^+\nabla_\mu\epsilon^{I\mu\nu}\alpha^{\alpha\beta}{}_I\nabla_\alpha A_\beta^- = \frac{1}{2}\int d^4x\sqrt{-g}\, A_\nu^+\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu H_-^{\dot{I}} \tag{5.164}$$

where we used the calculation above in the last equality. A last integration by parts gives the above formula.

Once we derive Maxwell equations for the fields, (5.162), we can obtain the corresponding first order equations for the potentials. Recall that ${}^\star F^{\mu\nu} = \alpha^{\mu\nu}{}_I H_+^I$, then Maxwell equations above imply $\nabla_\mu {}^+F^{\mu\nu} = 0$. As ${}^+F^{\mu\nu}$ is self-dual (hence the motivation for introducing the α matrices) then $\nabla_\mu {}^\star F_+{}^{\mu\nu} = 0$ as well, or equivalently $dF^+ = 0$. Thus we can solve Maxwell equations by introducing "another" potential, ${}^+F = dB$. This implies $H_i^+ = \frac{1}{4}\alpha_I^{\mu\nu\,+}F_{\mu\nu} = \frac{1}{2}\alpha_I^{\mu\nu}\nabla_\mu B_\nu$. Using (5.95) we can invert this relation

$$2\nabla_{[\mu}B_{\nu]} = {}^+F_{\mu\nu} = \alpha_{\mu\nu}{}^I H_I^+ = i\alpha_{\mu\nu}{}^I\epsilon_{I\alpha\beta}\nabla^\alpha A_+^\beta \tag{5.165}$$

Then,

$$2\epsilon^{I\mu\nu}\nabla_\mu B_\nu = i\epsilon^{I\mu\nu}\alpha_{\mu\nu}{}^J\epsilon_{J\alpha\beta}\nabla^\alpha A_+^\beta = 2\epsilon^{I\mu\nu}\nabla_\mu A_\nu^+ \tag{5.166}$$

where we noticed (5.113) and used (5.96)-(5.97). Recall that the potential A^+ is defined through a "curl" like this. Then, without loss of generality we can take $B \equiv A_+$, modulo gauge transformation.

Now, using both expressions for H_+:

$$i\epsilon^{I\mu\nu}\nabla_\mu A_\nu^+ = \frac{1}{2}\alpha^{\mu\nu I}\nabla_\mu A_\nu^+ \tag{5.167}$$

and using $2i\epsilon^{I\mu\nu} = \alpha^{\mu\nu I} - \bar{\alpha}^{\mu\nu I}$, we finally get the desired result:

$$\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu A_\nu^+ = 0 \tag{5.168}$$

F. Deriving the Noether current from the first-order action

In this section we derive the Noether current in first-order formalism by working directly with the standard variables A_+ and A_- and the action functional (5.47).

If the lagrangian density is the functional $L = L[A_+, A_-]$, then its variation under an infinitesimal duality rotation of the potential, $\delta A_\pm = \mp i\theta A_\pm$, yields

$$\begin{aligned}
\delta L &= \frac{\partial L}{\partial A_a^+}\delta A_a^+ + \frac{\partial L}{\partial \nabla_a A_b^+}\delta \nabla_a A_b^+ + c.c. \\
&= -\frac{1}{2}H_-^i \bar{\alpha}^{ab}{}_i(-i\theta)\nabla_a A_b^+ - \frac{1}{2}i\epsilon^{iab}\alpha^{cd}{}_i\nabla_c A_d^-(-i\theta)\nabla_a A_b^+ + c.c. \\
&= \frac{i\theta}{2}H_-^i \bar{\alpha}^{ab}{}_i\nabla_a A_b^+ + \frac{i\theta}{2}H_+^i \alpha^{cd}{}_i\nabla_c A_d^- + c.c. = 0 \equiv \nabla_a h^a \qquad (5.169)
\end{aligned}$$

We find that, unlike in second-order formalism, the duality rotation is an exact symmetry transformation of the first-order Lagrangian. The Noether current is now constructed as

$$\begin{aligned}
j^a &= \frac{\partial L}{\partial \nabla_a A_b^+}\delta A_b^+ + c.c. - h^a \qquad (5.170) \\
&= -\frac{1}{2}H_-^{\dot I}\bar{\alpha}^{\mu\nu}{}_{\dot I}(-i\theta)A_\nu^+ - \frac{1}{2}i\epsilon^{Iab}\alpha^{cd}{}_I\nabla_c A_d^-(-i\theta)A_b^+ + c.c. \\
&= \frac{i\theta}{2}[H_-^i \bar{\alpha}^{ab}{}_i A_b^+ - H_+^i \alpha^{ab}{}_i A_b^-] - \frac{\theta}{2}\epsilon^{Iab}A_b^+\alpha^{cd}{}_I\nabla_c A_d^- + c.c. \qquad (5.171)
\end{aligned}$$

Note: the last term vanishes on-shell. This term is precisely the one that makes this expression for the current different to that given by (5.64). Classically it does not matter, but perhaps quantum-mechanically could be important (actually something similar appears in the trace anomaly for Dirac fields, in which a piece in the stress energy tensor is proportional to the equations of motion and one neglects it because classically one is able to do it, but quantum fluctuations may be important there). However the charge agrees in both cases since the odd term above vanishes when contracted with n_a.

Let us evaluate now the divergence of the current:

$$\begin{aligned}
\nabla_a j^a &= \frac{i}{2}(\nabla_a H_-^{\dot I})\bar{\alpha}^{ab}{}_{\dot I}A_b^+ + \frac{i}{2}H_-^{\dot I}\bar{\alpha}^{ab}{}_{\dot I}\nabla_a A_b^+ - \frac{i}{2}(\nabla_a H_+^I)\alpha^{ab}{}_I A_b^- - \frac{i}{2}H_+^I \alpha^{ab}{}_I \nabla_a A_b^- \qquad (5.172) \\
&- \frac{1}{2}A_b^+\alpha^{cd}{}_I\nabla_a\epsilon^{Iab}\nabla_c A_d^- - \frac{1}{2}\epsilon^{Iab}\nabla_a A_b^+\alpha^{cd}{}_I\nabla_c A_d^- - \frac{1}{2}A_b^-\bar{\alpha}^{cd}{}_{\dot I}\nabla_a\epsilon^{\dot I ab}\nabla_c A_d^+ - \frac{1}{2}\epsilon^{\dot I ab}\nabla_a A_b^-\bar{\alpha}^{cd}{}_{\dot I}\nabla_c A_d^+
\end{aligned}$$

note that the second and last term cancel each other, as well as the fourth and the sixth terms. We are left with

$$\begin{aligned}
\nabla_a j^a &= \mathrm{Re}\left[i(\nabla_a H_-^{\dot I})\bar{\alpha}^{ab}{}_{\dot I}A_b^+\right] - \mathrm{Re}\left[A_b^+\alpha^{cd}{}_I\nabla_a\epsilon^{Iab}\nabla_c A_d^-\right] \\
&= \mathrm{Re}\left[i(\nabla_a F_-^{ab})A_b^+\right] + \mathrm{Re}\left[A_b^+ i\bar{\alpha}^{ab}{}_{\dot I}\nabla_a H_-^{\dot I}\right] = 2\mathrm{Re}\left[i(\nabla_a F_-^{ab})A_b^+\right] \qquad (5.173)
\end{aligned}$$

Now we would like to compare to the familiar expression that we got in second-order or lagrangian formalism (5.8). We have to be careful at this point. The relation between the 3-dimensional vector fields H^i_+ and potential A_+ is

$$H_I^+ = \frac{1}{2}\alpha^{ab}{}_I \nabla_a A_b^+ \tag{5.174}$$

valid only on-shell. This is equivalent to the equality ${}^+F = dA^+$. To pass from first to second order one has to do the Legendre transform, which amounts in the relation $E_i = \dot{A}_i$. It can be readily seen that the real part of the equation above gives this, while the imaginary part gives the wave equation for the potential $B^i = \epsilon^{ijk}\partial_j A_k = \dot{Z}^i$. By focusing only on the real part, i.e. $Re({}^+F) = Re(dA^+)$, it turns out that $F = dA$ and this leads to $\nabla_\mu {}^\star F^{\mu\nu} = 0$:

$$\nabla_a j_D^a \;\; = \;\; -(\nabla_a F^{ab}) Z_b \tag{5.175}$$

which is what we expected.

Appendix G. Lorentz Gauge fixing and the extension of the α's

As argued in the paragraphs following (5.47), there is a very nice way of incorporating the Lorentz gauge fixing term in the action into the language of α. This is analogous to what is usually done in the fermion case, in which besides the Pauli matrices one introduces a fourth matrix $\sigma^0 \equiv \mathbb{I}_{2\times 2}$ that conmutes with the others. In this appendix we shall check that the standard source-free Maxwell equations are still recovered with this additional input to the action. In this appendix $I = 1,2,3$ still.

Let us denote the "bare" action by

$$S_0[A_+, A_-] = -\frac{1}{2}\int d\Sigma_4 \left[H_-^{\dot{I}}\, \bar{\alpha}^{\mu\nu}{}_{\dot{I}} \nabla_\mu A_\nu^+ + H_+^I\, \alpha^{\mu\nu}{}_I \nabla_\mu A_\nu^- \right] \tag{5.176}$$

As shown in Appendix X variation with respect to A_μ^+ recover Maxwell equations $\alpha_I^{\mu\nu} \nabla_\nu H_+^I = 0$. However the extended action reads now

$$S[A_+, A_-] = S_0[A_+, A_-] + \frac{1}{2}\int d\Sigma_4 \left[H_-^0 \nabla_a A_+^a + H_+^0 \nabla_a A_-^a \right] \tag{5.177}$$

Variation with respecto to $H_\pm^0$ provides the Lorentz gauge fixing: $\nabla_a A_\pm^a = 0$. On the other hand, variation with respecto to A_ν^+ yields

$$0 = \frac{\delta S_M}{\delta A_\nu^+} = \bar{\alpha}_{\dot{I}}^{\mu\nu} \nabla_\mu H_-^{\dot{I}} - \frac{1}{2}\nabla^\nu H_-^0 \tag{5.178}$$

Now, on the one hand we can multiply this by t_ν, which taking into account that $D_{\dot{I}}H^{\dot{I}}_{-} = 0$ by construction (since $H^{\dot{I}}_{-} = -i\epsilon^{\dot{I}\mu\nu}\nabla_\mu A^{-}_{\nu}$ — see (5.139)), provides $t^\mu\nabla_\mu H^0_- \equiv \partial_t H^0_- = 0$. On the other hand, multiplying by ∇_ν one gets $\Box H^0_- = 0$ because of the identity $\bar{\alpha}^{\mu\nu}_{\dot{I}}\nabla_\nu\nabla_\mu H^{\dot{I}}_{-} = 0$. This last formula can be checked as follows:

$$\begin{aligned}\bar{\alpha}^{\mu\nu}_{\dot{I}}\nabla_\nu\nabla_\mu H^{\dot{I}}_{-} &= \nabla_\mu\nabla_\nu\bar{\alpha}^{\mu\nu}_{\dot{I}}H^{\dot{I}}_{-} = \nabla_\mu\nabla_\nu F^{\mu\nu}_{-} = \frac{1}{2}R_{\mu\nu\alpha}{}^{\mu}F^{\alpha\nu}_{-} + \frac{1}{2}R_{\mu\nu\alpha}{}^{\nu}F^{\mu\alpha}_{-} \\ &= \frac{1}{2}R_{\nu\alpha}F^{\alpha\nu}_{-} - \frac{1}{2}R_{\mu\alpha}F^{\mu\alpha}_{-} = 0 \end{aligned} \tag{5.179}$$

If both $\Box H^0_- = 0$ and $\partial_t H^0_- = 0$ hold, then $D_I D^I H^0_- = 0$ holds. Choosing appropriate boundary conditions one can solve the equation and get H^0_- be constant. Turning back to (5.178), $\nabla_\mu H^0_- = 0$ one gets the desired source-free Maxwell equations. An identical reasoning can be applied for H_+. Following now the same arguments as in Appendix E one gets to the extended first-order equations of motion for the potentials:

$$\bar{\alpha}^{\mu\nu}_{\dot{I}'}\nabla_\mu A^{+}_{\nu} = 0 \tag{5.180}$$

where $I = 0, 1, 2, 3$ now.

Appendix H. Definition of Ψ and β^μ and properties

In section (5.4) we wrote the action functional as similar to the Dirac case. In the successive steps of reasoning we introduced without much explanation certain elements that were used throughout the remaining text. We define here those elements in detail and provide their main properties.

Given the complex potentials $A^{\pm}_{\mu}$ and the electromagnetic self- and antiself- dual fields, H^I_+ and $H^{\dot{I}}_{-}$, we can construct the direct sum of their correspondings space of functions. Let X denote this space. An arbitrary element of this space will have the form:

$$\Psi = \begin{pmatrix} A^{+}_{\mu} \\ H^{I}_{+} \\ A^{-}_{\mu} \\ H^{\dot{I}}_{-} \end{pmatrix} \in X \tag{5.181}$$

We define now the vector-valued mapping $\beta^\mu : X \longrightarrow X$ as

$$\beta^\mu\Psi = i\begin{pmatrix} \bar{\alpha}^{\mu\nu}_{\dot{I}}H^{\dot{I}}_{-} \\ -\alpha^{\mu\nu}_{I}A^{-}_{\nu} \\ \alpha^{\mu\nu}_{I}H^{I}_{+} \\ -\bar{\alpha}^{\mu\nu}_{\dot{I}}A^{+}_{\nu} \end{pmatrix} \tag{5.182}$$

which is well-defined ($\beta^\mu \Psi \in X, \forall \Psi \in X$) and is linear. This allows to write

$$\beta^\mu = -i \begin{pmatrix} 0 & 0 & 0 & -\bar{\alpha}^{\mu\nu}_{\dot{I}} \\ 0 & 0 & \alpha^{\mu\nu}_{I} & 0 \\ 0 & -\alpha^{\mu\nu}_{I} & 0 & 0 \\ \bar{\alpha}^{\mu\nu}_{\dot{I}} & 0 & 0 & 0 \end{pmatrix} \tag{5.183}$$

Define now the product of two β^μ as the composite operation, $\beta^\mu\beta^\nu : X \longrightarrow X$, by $(\beta^\mu\beta^\nu)\Psi = \beta^\mu(\beta^\nu\Psi)$, that is well-defined and linear. This leads to

$$\beta^\nu\beta^\mu = \begin{pmatrix} \bar{\alpha}^{\nu}{}_{\nu\dot{J}}\bar{\alpha}^{\mu}_{\rho}{}^{\dot{J}} & 0 & 0 & 0 \\ 0 & \alpha^{\nu\rho}{}_{J}\alpha^{\mu}{}_{\rho I} & 0 & 0 \\ 0 & 0 & \alpha^{\nu}{}_{\nu J}\alpha^{\mu}_{\rho}{}^{J} & 0 \\ 0 & 0 & 0 & \bar{\alpha}^{\nu\rho}{}_{\dot{J}}\bar{\alpha}^{\mu}{}_{\rho\dot{I}} \end{pmatrix} \tag{5.184}$$

whose symmetric and antisymmetric parts, using (5.103)-(5.104), recover formulas (5.55).

The "chiral" operation is a linear application $\beta_5 : X \longrightarrow X$ defined by formula (5.56). It yields

$$\begin{aligned}
\beta_5 &= \frac{i}{4!}\epsilon_{abcd}\beta^a\beta^b\beta^c\beta^d = \frac{i}{4!}\epsilon_{abcd}\beta^{[a}\beta^{b]}\beta^{[c}\beta^{d]} \\
&= \frac{i}{6}\epsilon_{abcd} \begin{pmatrix} [{}^+\Sigma^{ab}]_{\sigma\rho}[{}^+\Sigma^{cd}]^{\rho}{}_{\alpha} & 0 & 0 & 0 \\ 0 & [{}^+\Sigma^{ab}]_{IK}[{}^+\Sigma^{cd}]^{K}{}_{J} & 0 & 0 \\ 0 & 0 & [{}^-\Sigma^{ab}]_{\sigma\rho}[{}^-\Sigma^{cd}]^{\rho}{}_{\alpha} & 0 \\ 0 & 0 & 0 & [{}^-\Sigma^{ab}]_{\dot{I}\dot{K}}[{}^-\Sigma^{cd}]^{\dot{K}}{}_{\dot{J}} \end{pmatrix} \\
&= \frac{1}{3} \begin{pmatrix} [{}^+\Sigma_{cd}]_{\sigma\rho}[{}^+\Sigma^{cd}]^{\rho}{}_{\alpha} & 0 & 0 & 0 \\ 0 & [{}^+\Sigma_{cd}]_{IK}[{}^+\Sigma^{cd}]^{K}{}_{J} & 0 & 0 \\ 0 & 0 & -[{}^-\Sigma_{cd}]_{\sigma\rho}[{}^-\Sigma^{cd}]^{\rho}{}_{\alpha} & 0 \\ 0 & 0 & 0 & -[{}^-\Sigma_{cd}]_{\dot{I}\dot{K}}[{}^-\Sigma^{cd}]^{\dot{K}}{}_{\dot{J}} \end{pmatrix} \\
&= \frac{1}{3} \begin{pmatrix} -2[{}^+\Sigma_{\sigma\rho}]^{\rho}{}_{\alpha} & 0 & 0 & 0 \\ 0 & -2[{}^+\Sigma_{IK}]^{K}{}_{J} & 0 & 0 \\ 0 & 0 & 2[{}^-\Sigma_{\sigma\rho}]^{\rho}{}_{\alpha} & 0 \\ 0 & 0 & 0 & 2[{}^-\Sigma_{\dot{I}\dot{K}}]^{\dot{K}}{}_{\dot{J}} \end{pmatrix} \\
&= \begin{pmatrix} -g_{\mu\nu} & 0 & 0 & 0 \\ 0 & -\eta_{IJ} & 0 & 0 \\ 0 & 0 & g_{\mu\nu} & 0 \\ 0 & 0 & 0 & \eta_{\dot{I}\dot{J}} \end{pmatrix}
\end{aligned} \tag{5.185}$$

and satisfies

$$\beta_5^2 = \begin{pmatrix} g_{\mu\nu} & 0 & 0 & 0 \\ 0 & \eta_{IJ} & 0 & 0 \\ 0 & 0 & g_{\mu\nu} & 0 \\ 0 & 0 & 0 & \eta_{\dot{I}\dot{J}} \end{pmatrix}, \qquad \{\beta_5, \beta^\mu\} = 0 \tag{5.186}$$

Note that both $\beta_5\beta^\mu : X \longrightarrow X$ and $\beta^\mu\beta_5 : X \longrightarrow X$ are well-defined as composite operations. The duality transformation is defined by means of a linear application $T_\theta : X \longrightarrow X$, $T_\theta = e^{i\theta\beta_5}$, $\theta \in \mathbb{R}$.

Let X^* be now the dual space, the space of bounded linear functionals $\phi : X \longrightarrow \mathbb{C}$. We shall work on a particular class of elements of this space: given $\Psi \in X$, we shall define $\bar{\Psi} \in X^*$ by

$$\bar{\Psi} := \begin{pmatrix} A_\mu^+ & H_+^I & A_\mu^- & H_-^{\dot{I}} \end{pmatrix} \tag{5.187}$$

This way, the action functional (5.53) is a well-defined quantity.

Lastly, we introduce the "complex-conjugation" in this space of elements by $\beta : X \longrightarrow X$,

$$\beta\Psi = \begin{pmatrix} A_\mu^- \\ H_-^I \\ A_\mu^+ \\ H_+^{\dot{I}} \end{pmatrix} \tag{5.188}$$

which leads to

$$\beta = \begin{pmatrix} 0 & 0 & \delta_\nu^\mu & 0 \\ 0 & 0 & 0 & \delta_{\dot{I}}^I \\ \delta_\nu^\mu & 0 & 0 & 0 \\ 0 & \delta_I^{\dot{I}} & 0 & 0 \end{pmatrix} \tag{5.189}$$

This is useful since now $\bar{\Psi}\beta\Psi \in \mathbb{C}$ is well-defined, and moreover $\bar{\Psi}\beta\Psi \geq 0$. Thus, the inner product used in 5.5.2 is well-defined.

Appendix I. Explicit calculations of the electromagnetic duality anomaly

From (5.69) and (5.70) and text in between we can write

$$\langle \nabla_\mu j_D^\mu \rangle_{ren} = \frac{\hbar}{32\pi^2} \lim_{m\to 0} m \sum_{k=0}^{2} \mathrm{Tr}\left[(i\beta^\mu \nabla_\mu + m)\beta_5 E_k(x) \right] \int_0^\infty d(i\tau) e^{-i\tau m^2} (i\tau)^{(k-2)} \tag{5.190}$$

In order to make the integral convergent we need to add here a tiny imaginary contribution to the mass, which amounts in a choice for the Green function. It corresponds to $m^2 \to m^2 - i\epsilon$. Solving the integral yields

$$\langle \nabla_\mu j^\mu_D \rangle_{ren} = \frac{\hbar}{32\pi^2} \lim_{m\to 0} m \sum_{k=0}^{2} \Gamma(k-1)(im^2)^{1-k} \mathrm{Tr}\left[(i\beta^\mu \nabla_\mu + m)\beta_5 E_k(x)\right] \quad (5.191)$$

In order to obtain the Heat Kernel coefficients we need to obtain a quadratic equation of motion for the field. Since $\nabla_\mu \beta^\nu(x) = 0$, the quadratic operator is $0 = \beta^\mu\beta^\nu\nabla_\mu\nabla_\nu\Psi = \left[\beta^{(\mu}\beta^{\nu)} + \beta^{[\mu}\beta^{\nu]}\right]\nabla_\mu\nabla_\nu\Psi$

$$\{\mathbb{I} g^{\mu\nu}\nabla_\mu\nabla_\nu - \begin{pmatrix} {}^+\Sigma^{\mu\nu} & 0 & 0 & 0 \\ 0 & {}^+\Sigma^{\mu\nu} & 0 & 0 \\ 0 & 0 & {}^-\Sigma^{\mu\nu} & 0 \\ 0 & 0 & 0 & {}^-\Sigma^{\mu\nu} \end{pmatrix} [\nabla_\mu, \nabla_\nu]\}\Psi = 0$$

If we express $(\Box + \mathcal{Q})\Psi = 0$, then

$$\mathcal{Q} \equiv -\frac{1}{2} R_{\mu\nu\alpha\beta} \begin{pmatrix} {}^+\Sigma^{\mu\nu}\Sigma^{\alpha\beta} & 0 & 0 & 0 \\ 0 & 2\,{}^+\Sigma^{\mu\nu}\,{}^+\Sigma^{\alpha\beta} & 0 & 0 \\ 0 & 0 & {}^-\Sigma^{\mu\nu}\Sigma^{\alpha\beta} & 0 \\ 0 & 0 & 0 & 2\,{}^-\Sigma^{\mu\nu}\,{}^-\Sigma^{\alpha\beta} \end{pmatrix} \quad (5.192)$$

Note: given $\nabla_\mu\Psi = \partial_\mu\Psi - \frac{1}{2}\omega_\mu{}^{ab} M_{ab}\Psi$, with M_{ab} a generator of the Lorentz symmetry group, it can be checked that $\left[\nabla_\mu, \nabla_\nu\right]\Psi = \frac{1}{2}R_{\mu\nu}{}^{ab}M_{ab}\Psi$. The proof can be found in Parker-Toms, pages 225-226. See Formulas (5.270)-(5.271).

Since $E_0 = \mathbb{I}$, then $\mathrm{Tr}(\beta_5 E_0) = 0$ and $\mathrm{Tr}(\beta^\mu\beta_5 E_0) = 0$. On the other hand, $E_1 = \frac{1}{6}R\mathbb{I} - \mathcal{Q}$, and consequently $\mathrm{Tr}(\beta_5 E_1) = -\mathrm{Tr}(\beta_5\mathcal{Q})$,

$$\begin{aligned} \mathrm{Tr}(\beta_5 E_1) &= iR_{\mu\nu\alpha\beta} Im\, \mathrm{Tr}\{{}^+\Sigma^{\mu\nu}(2\,{}^+\Sigma^{\alpha\beta} + \Sigma^{\alpha\beta})\} = -\frac{5}{4}R_{\mu\nu\alpha\beta}\mathrm{Tr}\{\Sigma^{\mu\nu\star}\Sigma^{\alpha\beta}\} \\ &= -\frac{5}{4}R_{\mu\nu\alpha\beta}\Sigma^{\mu\nu}_{IJ}{}^{\star}\Sigma^{\alpha\beta JI} = -\frac{5}{2}R_{IJ\alpha\beta}\epsilon^{IJ\alpha\beta} = -\frac{5}{2}R_{\mu\nu\alpha\beta}\epsilon^{\mu\nu\alpha\beta} = 0 \quad (5.193) \end{aligned}$$

where in the last step we used the Bianchi identity. The equality $\mathrm{Tr}(\beta^\mu\beta_5 E_1) = 0$ can be easily obtained by noting that no diagonal terms appear in the "matrix" $\beta^\mu\beta_5 E_1$. The same argument holds for $\mathrm{Tr}(\beta^\mu\beta_5 E_2) = 0$. After all, (5.191) reduces to

$$\begin{aligned} \langle \nabla_\mu j^\mu_D \rangle &= \frac{-i\hbar}{32\pi^2}\mathrm{Tr}(\beta_5 E_2) \quad (5.194) \\ &= \frac{-i\hbar}{32\pi^2}\left[\frac{1}{12}\mathrm{Tr}(\beta_5 W_{\mu\nu}W^{\mu\nu}) + \frac{1}{2}\mathrm{Tr}(\beta_5\mathcal{Q}^2)\right] \end{aligned}$$

Recall that

$$W_{\mu\nu}\Psi = [\nabla_\mu, \nabla_\nu]\Psi = \frac{1}{2}R_{\mu\nu\alpha\beta}\begin{pmatrix} \Sigma^{\alpha\beta} & 0 & 0 & 0 \\ 0 & 2\,{}^{+}\Sigma^{\alpha\beta} & 0 & 0 \\ 0 & 0 & \Sigma^{\alpha\beta} & 0 \\ 0 & 0 & 0 & 2\,{}^{-}\Sigma^{\alpha\beta} \end{pmatrix}\Psi$$

and hence

$$W_{\mu\nu}W^{\mu\nu} = \frac{1}{4}R_{\mu\nu\alpha\beta}R^{\mu\nu}{}_{\sigma\rho}\begin{pmatrix} \Sigma^{\alpha\beta}\Sigma^{\sigma\rho} & 0 & 0 & 0 \\ 0 & 4\,{}^{+}\Sigma^{\alpha\beta\,+}\Sigma^{\sigma\rho} & 0 & 0 \\ 0 & 0 & \Sigma^{\alpha\beta}\Sigma^{\sigma\rho} & 0 \\ 0 & 0 & 0 & 4\,{}^{-}\Sigma^{\alpha\beta\,-}\Sigma^{\sigma\rho} \end{pmatrix}$$

so that

$$\begin{aligned}
\mathrm{Tr}(\beta_5 W_{\mu\nu}W^{\mu\nu}) &= -\frac{1}{4}R_{\mu\nu\alpha\beta}R^{\mu\nu}{}_{\sigma\rho}2i\,\mathrm{Im}\,\mathrm{Tr}\left[\Sigma^{\alpha\beta}\Sigma^{\sigma\rho} + 4\,{}^{+}\Sigma^{\alpha\beta\,+}\Sigma^{\sigma\rho}\right] \\
&= -2iR_{\mu\nu\alpha\beta}R^{\mu\nu}{}_{\sigma\rho}\mathrm{Im}\,\mathrm{Tr}\left[{}^{+}\Sigma^{\alpha\beta\,+}\Sigma^{\sigma\rho}\right] \\
&= -\frac{i}{2}R_{\mu\nu\alpha\beta}R^{\mu\nu}{}_{\sigma\rho}\mathrm{Tr}\left[\Sigma^{\alpha\beta\,\star}\Sigma^{\sigma\rho} + {}^{\star}\Sigma^{\alpha\beta}\Sigma^{\sigma\rho}\right] \\
&= -iR_{\mu\nu\alpha\beta}R^{\mu\nu}{}_{\sigma\rho}\mathrm{Tr}\left[\Sigma^{\alpha\beta\,\star}\Sigma^{\sigma\rho}\right] \\
&= 2iR_{\mu\nu IJ}R^{\mu\nu}{}_{\sigma\rho}\epsilon^{\sigma\rho IJ} = -2iR_{\mu\nu\alpha\beta}R^{\mu\nu}{}_{\sigma\rho}\epsilon^{\sigma\rho\alpha\beta} \\
&= 4iR_{\mu\nu\alpha\beta}{}^{\star}R^{\mu\nu\alpha\beta}
\end{aligned} \tag{5.195}$$

On the other hand,

$$\begin{aligned}
\mathrm{Tr}(\beta_5\mathscr{Q}^2) &= -\frac{1}{4}R_{\mu\nu\alpha\beta}R_{abcd}\,2i\,\mathrm{Im}\,\mathrm{Tr}\{{}^{+}\Sigma^{\mu\nu}\Sigma^{\alpha\beta\,+}\Sigma^{ab}\Sigma^{cd} + 4\,{}^{+}\Sigma^{\mu\nu\,+}\Sigma^{\alpha\beta\,+}\Sigma^{ab\,+}\Sigma^{cd}\} \\
&= -\frac{i}{8}R_{\mu\nu\alpha\beta}R_{abcd}\mathrm{Tr}\{\Sigma^{\mu\nu}\Sigma^{\alpha\beta\,\star}\Sigma^{ab}\Sigma^{cd} + {}^{\star}\Sigma^{\mu\nu}\Sigma^{\alpha\beta}\Sigma^{ab}\Sigma^{cd} \\
&+ {}^{\star}\Sigma^{\mu\nu}\Sigma^{\alpha\beta}\Sigma^{ab}\Sigma^{cd} + \Sigma^{\mu\nu\,\star}\Sigma^{\alpha\beta}\Sigma^{ab}\Sigma^{cd} \\
&+ \Sigma^{\mu\nu}\Sigma^{\alpha\beta\,\star}\Sigma^{ab}\Sigma^{cd} + \Sigma^{\mu\nu}\Sigma^{\alpha\beta}\Sigma^{ab\,\star}\Sigma^{cd} \\
&+ {}^{\star}\Sigma^{\mu\nu\,\star}\Sigma^{\alpha\beta\,\star}\Sigma^{ab}\Sigma^{cd} + {}^{\star}\Sigma^{\mu\nu\,\star}\Sigma^{\alpha\beta}\Sigma^{ab\,\star}\Sigma^{cd} \\
&+ {}^{\star}\Sigma^{\mu\nu}\Sigma^{\alpha\beta\,\star}\Sigma^{ab\,\star}\Sigma^{cd} + \Sigma^{\mu\nu\,\star}\Sigma^{\alpha\beta\,\star}\Sigma^{ab\,\star}\Sigma^{cd}\}
\end{aligned} \tag{5.196}$$

Using the Bianchi identity (${}^{\star}R^{a}{}_{bac} = 0$), each of these terms vanishes. Finally then,

$$\begin{aligned}
\langle\nabla_\mu j^\mu_D\rangle &= \frac{-\hbar i}{32\pi^2}\left[\frac{1}{12}\mathrm{Tr}(\beta_5 W_{\mu\nu}W^{\mu\nu}) + \frac{1}{2}\mathrm{Tr}(\beta_5\mathscr{Q}^2)\right] \\
&= \frac{-\hbar}{96\pi^2}R^{\mu\nu\alpha\beta}[{}^{\star}R]_{\mu\nu\alpha\beta}.
\end{aligned} \tag{5.197}$$

Since the heat kernel asymptotic series of the heat kernel does not depend on the vacuum state chosen, this expectation value is (vacuum) state independent.

Appendix J. Trace anomaly

As a non trivial check, one might be interested in reproducing the trace anomaly using the Dirac-type approach introduced in this work. The calculation is straightforward following the recipes given above, but considerably more lenghty since there are non-vanishing contributions coming from all terms in the heat kernel expansion. We can predict, though, what the result should read by following Chirstensen and Duff analysis [61, 62].

Our formulation would yield a trace anomaly determined by the $A_{\pm}$, $H_{\pm}$ fields, which transform under Lorentz trasformations according to the (1/2,1/2) and (1,0)/(0,1) representations. From the action (5.53) one can conclude that:

$$\begin{aligned} \langle T^{\mu}_{\mu} \rangle &= -\frac{1}{4}\langle \bar{\Psi} D \Psi \rangle + 2\langle c \Box c \rangle \\ &= -\frac{1}{4}[b_4(1/2,1/2) + b_4(1,0) + b_4(1/2,1/2) + b_4(0,1)] - 2b_4(0,0) \end{aligned} \tag{5.198}$$

where c represents one of the two ghosts (we omitted them along the text since they are irrelevant for the duality anomaly), that is described by a minimally coupled, massless, scalar field. In the second line we introduced $b_4(A,B)$; it denotes the coefficient in the heat kernel expansion of 4th order in derivatives of the metric, and can be determined by a knowledge of the associated Lorentz representation upon which the field of interest transforms with, as calculated in [61, 62]. The $-\frac{1}{4}$ factor would come from the weird normalization in the action (5.53). It can be deduced from [61] that $b_4(1,0) + b_4(0,1) = -6b_4(1/2,1/2)$. Then we recover

$$\langle T^{\mu}_{\mu} \rangle = b_4(1/2,1/2) - 2b_4(0,0) \tag{5.199}$$

which is the standard result [48].

Chapter 6

Conclusions and outlook

This thesis is the result of the research carried out by the author during the last 5 years in collaboration with his supervisors and other people. The common topic is the quantum theory of fields propagating in curved spacetimes. Although more 50 years have passed since the first quantitative paper on the subject [147], the topic is still of great importance nowadays, not only because it continues to give new insights in the foundations of a supposed quantum theory of gravity, but it also leads to new phenomenological implications that, today more than ever, could be tested in experiments. To the present day, cosmological missions have reached such a level of sensitivity that quantum effects of fields during the early universe must be taken into account for the correct interpretation of both CMB and large scale structure data analysis. On the other hand, the detection of gravitational waves by the LIGO-Virgo collaboration, together with the electromagnetic counterpart measured by multimessenger astronomy, has opened up a new era in astrophysics. There is no doubt that this rapidly increasing research area will provide considerable amounts of data from astrophysical sources that will allow a better understanding, not only of general relativity, but also of the consequences of the quantum theory of fields around black holes scenarios.

With this motivation in mind, in the first part of this thesis we have dealt with several questions related to primordial cosmology. In Chapter 2 we analyzed the most important observable supporting the theory of inflation, the angular power spectrum, from a different, not conventional, perspective. We reexamined this quantity from a spacetime point of view and found that the standard result considered in the literature actually diverges in the ultraviolet. This is complementary to the work done in Fourier space a decade ago. The usually referred assumption of taking the large distance limit in the two-point function of primordial cosmological perturbations, leading to the well-known – and experimentally checked – scale invariance of the power spectrum, serves as a natural regulator of the UV divergences. This, however, turns out to correspond to a zeroth-order adiabatic subtraction (as it is known in the specialized literature of renormalization). We argued that subtractions up to second adiabatic order should be more natural from a physical point of view. By doing this one gets an additional (scale invariant) term in the power spectrum (see (2.39) for the final result), and its possible observable consequences were discussed.

Chapter 3 is a throughout text that deals with the calculation of the renormalized stress-energy tensor of spin 1/2 fields in expanding universes. This was achieved both when the fermion field propagates freely and when the field is interacting with an external scalar field through a Yukawa coupling. Final expressions for the renormalized values can be looked up in (3.72) and (3.73); and (3.157), (3.158), and (3.171). The underlying motivation was finding the possibility of carrying out the study of matter (namely, the production of particles and energy) not only during the inflationary regime, but also during the preheating era, when all matter known in the universe is supposed to have originated from the inflaton field oscillations. Some examples of cosmological interest using this formalism were discussed using analytical approximations (see section 3.4 and Appendix B) and some using numerical approaches (see Appendix E). The work done here paves the way for future projects regarding numerical implementations of matter particle production during the early universe.

The cosmology study ends with Chapter 4. In single field models of inflation, it is normally assumed that the only field relevant for the computation of physical observables is the one that drives the exponential expansion itself. In this chapter we considered the presence and influence of a large number N of test or spectator light scalar fields, fields that do not couple directly to the inflaton field, only propagate in the spacetime background. This is of interest since some fundamental theories (i.e. string theory, supergravity, etc) generally require the existence of large amounts of light fields for self-consistency, and their presence should certainly have some imprint somewhere, for instance in CMB observables. It turns out that they amount to a running contribution ($\log k$ dependence) to both the scalar and tensorial Fourier spectra. We argued that, though normally neglected in standard approaches to inflation, their impact in the tensor spectrum could be well relevant to put bounds on the possible number of N light scalar fields present in the universe. The bound obtained and associated discussion is found in (4.58) and ahead. The results could constrain phenomenological models that require such huge amounts of light fields to solve the hierarchy problem in the standard model of particles.

Turning then to more fundamental aspects of quantum field theory, in Chapter 5 we dealt with the electromagnetic duality transformation of source-free Maxwell theory in curved spacetimes. Classically, this was shown to be a dynamical symmetry of the Maxwell action if no electric charges of currents are present. However, motivated by the knowledge of other anomalies, the natural question of whether this symmetry could be extended to the quantum theory was posed and analyzed in detail. Indeed, as first shown in 1969 [10, 44, 122], this is a non-trivial issue: classical symmetries may fail to hold in the quantum regime due to off-shell contributions, coming from quantum corrections. In this chapter we calculated the vacuum expectation value of the divergence of the Noether current, and found that it is not vanishing due to the renormalization subtractions that one needs to consider in order to properly account for the UV divergences. The particular value obtained can be looked up in (5.77). This result thus leads to an anomaly in the quantum theory. We also checked the calculation by following Fujikawa's interpretation: the measure of the path integral — this latter one yielding the transition amplitudes between quantum states of the field in different times — transforms non-trivially, providing the anomaly. Finally, we ended the chapter by commenting on potential physical implications of this in astrophysics and gravitational wave physics (see section 5.6).

Concerning this last point, our results opens up several avenues for future research, some of which I proceed to describe now. Immediate questions arise: how should we interpret the electromagnetic duality anomaly from a physical point of view? Does it mean that an initial sample of photons get polarized if they pass through a strong gravitational background? Moreover, do deviations from general relativity predict distinctive signatures? Are there observable implications of all this, say, in astrophysics? Does this anomaly arise in gravitational radiation as well? We plan to address these questions in the near future by working out the specific research directions listed below. Answering them will help us to understand the implications of the anomaly from a measurable viewpoint, and this will be useful in testing either the validity of general relativity or quantum field theory in curved spacetimes in an astrophysical context.

1. The detailed calculation of the electromagnetic duality anomaly was only carried out considering the vacuum expectation value of the duality current j^μ, i.e. we computed the quantity $\langle 0| \nabla_\mu j^\mu |0\rangle \neq 0$. This polarization effect could be interpreted as spontaneous (asymmetric) creation of particles from vacuum due to the strong gravitational dynamics, similarly to the celebrated Hawking effect. If, instead, the initial state describing the electromagnetic field had already a distribution of photons, we would expect an stimulated contribution to this phenomenon from the quantum state, which must be much more significant in any practical situation of astrophysical or cosmological interest. Our plan is to determine how big this effect can be. In other words, we need to calculate $\langle \rho \big| \nabla_\mu j^\mu \big| \rho \rangle$ in detail, where $\big|\rho\rangle$ represents a mixed state describing photons with different helicities, frecuencies, etc. Then, we have to figure out how this quantity is measured by or imprinted in a quantum device.

We shall borrow additional techniques from quantum field theory in curved spacetime and general relativity. The use of Unruh-DeWitt detectors, in particular, will be important to understand in more depth what physically the impact of this anomaly on photons could be, since these detectors can tell what a test field or observer is able to measure or not. Asympotic analysis in General Relativity will also be important to address the question of incoming radiation from past null infinity and the scattered radiation to future null infinity. On the other hand, techniques usually employed in quantum information science might be well relevant for analyzing the issue of the choice of quantum state, i.e. for instance by means of density operators.

2. An additional assumption in our computation of the electromagnetic duality anomaly was the use of conventional general relativity. Since the result is a purely geometrical contribution, it could be interesting to investigate whether modified theories of gravity provide new terms to $\langle 0| \nabla_\mu j^\mu |0\rangle$ or $\langle \rho \big| \nabla_\mu j^\mu \big| \rho \rangle$. If so, this could serve to predict deviations from general relativity in strong gravitational scenarios, for instance through new signatures in GW emission. Several approaches in field theory or differential geometry could be investigated here. To give an example, if we allow torsion or non-metricity on the connection, apart from curvature additional contributions to the above expectation value might arise; etc.

3. Gravitational waves also have two radiative degrees of freedom, namely the well-known h_+ and $h_\times$ modes, and there exists an analogue of electromagnetic duality symmetry. Consequently, at the linearized level the notion of a charge measuring the state of polarization of gravitational radiation should be available as well. We plan to examine this in detail and study whether there is an anomaly in the theory of gravitons propagating in a non-trivial curved spacetime, i.e. if the background distinguishes (quantum-mechanically) between the two degrees of freedom. This could be particularly interesting in future measurements of polarization of gravitational waves.

4. Our eventual plan would be to search signatures of all this in astrophysics, namely in black hole physics and gravitational wave emission. Working in this area will imply getting familiarized with astrophysical sources, such as binary mergers or dynamical black holes, as well as the analytical techniques involved to extract physical information (asymptotic analysis, black hole multipole moments, asymptotic charges, etc). Questions such as if the effect grows linearly with the mass of the system, or if the polarization or angular momentum carried away by gravitational waves plays any role, will be addressed. Close collaboration with numerical relativists is expected.

Bibliography

[1] *Einstein telescope*, http://www.et-gw.eu/.

[2] *The fermi gamma-ray space telescope*, https://fermi.gsfc.nasa.gov/science/.

[3] *The international gamma-ray astrophysics laboratory*, http://www.isdc.unige.ch/integral/.

[4] *The laser interferometer space antenna*, http://sci.esa.int/lisa/.

[5] *Ligo - virgo collaboration*, https://www.ligo.org/detections.php.

[6] *Planck collaboration*, https://www.cosmos.esa.int/web/planck/publications.

[7] *The sloan digital sky survey*, http://www.sdss.org/science/publications/.

[8] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards, New York: Dover Publications, 1972.

[9] P. A. R. Ade et al., *Planck 2015 results. XX. Constraints on inflation*, Astron. Astrophys., 594 (2016), p. A20.

[10] S. L. Adler, *Axial-vector vertex in spinor electrodynamics*, Phys. Rev., 177 (1969), pp. 2426–2438.

[11] P. Adshead, R. Easther, and E. A. Lim, *The 'in-in' Formalism and Cosmological Perturbations*, Phys. Rev., D80 (2009), p. 083521.

[12] I. Agullo, A. del Rio, and J. Navarro-Salas, *Classical and quantum aspects of electromagnetic duality transformations*, (To appear).

[13] I. Agullo, A. del Rio, and J. Navarro-Salas, *Electromagnetic duality anomaly in curved spacetimes*, Phys. Rev. Lett., 118 (2017), p. 111301.

[14] I. AGULLO, A. DEL RIO, AND J. NAVARRO-SALAS, *Gravity and handedness of photons*, International Journal of Modern Physics D, 26 (2017), p. 1742001.

[15] I. AGULLO, A. LANDETE, AND J. NAVARRO-SALAS, *Electric-magnetic duality and renormalization in curved spacetimes*, Phys.Rev.D, (2014), p. 124067.

[16] I. AGULLO, J. NAVARRO-SALAS, G. OLMO, AND L. PARKER, *Inflation, quantum fields, and cmb anisotropies*, Gen. Relativ. Gravit., (2009), p. 2301.

[17] I. AGULLÓ, J. NAVARRO-SALAS, G. J. OLMO, AND L. PARKER, *Revising the predictions of inflation for the cosmic microwave background anisotropies*, Phys. Rev. Lett., 103 (2009), p. 061301.

[18] I. AGULLO, J. NAVARRO-SALAS, G. J. OLMO, AND L. PARKER, *Revising the observable consequences of slow-roll inflation*, Phys. Rev. D, 81 (2010), p. 043514.

[19] I. AGULLO, J. NAVARRO-SALAS, G. J. OLMO, AND L. PARKER, *Remarks on the renormalization of primordial cosmological perturbations*, Phys. Rev. D, 84 (2011), p. 107304.

[20] M. ALCUBIERRE, *Introduction to 3+1 numerical relativity*, International series of monographs on physics, Oxford Univ. Press, Oxford, 2008.

[21] A. L. ALINEA, *Adiabatic regularization of power spectra in nonminimally coupled chaotic inflation*, Journal of Cosmology and Astroparticle Physics, 2016 (2016), p. 027.

[22] A. L. ALINEA, T. KUBOTA, Y. NAKANISHI, AND W. NAYLOR, *Adiabatic regularisation of power spectra in k -inflation*, Journal of Cosmology and Astroparticle Physics, 2015 (2015), p. 019.

[23] B. ALLEN, *Vacuum states in de sitter space*, Phys. Rev. D, (1985), p. 3136.

[24] B. ALLEN AND A. FOLACCI, *Massless minimally coupled scalar field in de sitter space*, Phys. Rev. D, (1987), p. 3771.

[25] P. ANDERSON AND L. PARKER, *Adiabatic regularization in closed robertson-walker universes*, Phys. Rev. D, (1987), p. 2963.

[26] P. R. ANDERSON, *Effects of Quantum Fields on Singularities and Particle Horizons in the Early Universe. 3. The Conformally Coupled Massive Scalar Field*, Phys. Rev., D32 (1985), p. 1302.

[27] P. R. ANDERSON AND W. EAKER, *Analytic approximation and an improved method for computing the stress-energy of quantized scalar fields in robertson-walker spacetimes*, Phys. Rev. D, 61 (1999), p. 024003.

[28] P. R. ANDERSON, C. MOLINA-PARS, D. EVANICH, AND G. B. COOK, *Study of the preheating phase of chaotic inflation*, Phys. Rev. D, 78 (2008), p. 083514.

[29] N. ARKANI-HAMED, T. COHEN, R. T. D'AGNOLO, A. HOOK, H. D. KIM, AND D. PINNER, *Solving the Hierarchy Problem at Reheating with a Large Number of Degrees of Freedom*, Phys. Rev. Lett., 117 (2016), p. 251801.

[30] A. ASHTEKAR, *Asymptotic quantization: based on 1984 Naples lectures*, Monographs and textbooks in physical science, Bibliopolis, 1987.

[31] A. ASHTEKAR AND B. BONGA, *On the ambiguity in the notion of transverse traceless modes of gravitational waves*, Gen. Rel. Grav., 49 (2017), p. 122.

[32] A. ASHTEKAR AND B. KRISHNAN, *Isolated and dynamical horizons and their applications*, Living Rev. Rel., 7 (2004), p. 10.

[33] A. ASHTEKAR AND R. TATE, *Lectures on Non-perturbative Canonical Gravity*, Advanced series in astrophysics and cosmology, World Scientific, 1991.

[34] I. AVRAMIDI, *A covariant technique for the calculation of the one-loop effective action*, Nucl. Phys. B, (1991), p. 712.

[35] I. G. AVRAMIDI, *The covariant technique for the calculation of the heat kernel asymptotic expansion*, Phys. Lett. B, (1990), p. 92.

[36] J. BAACKE, K. HEITMANN, AND C. PÄTZOLD, *Nonequilibrium dynamics of fermions in a spatially homogeneous scalar background field*, Phys. Rev. D, 58 (1998), p. 125013.

[37] J. BAACKE AND C. PÄTZOLD, *Renormalization of the nonequilibrium dynamics of fermions in a flat frw universe*, Phys. Rev. D, 62 (2000), p. 084008.

[38] J. M. BARDEEN, P. J. STEINDHART, AND M. S. TURNER, *Spontaneous creation of almost scale-free density perturbations in an inflationary universe*, Phys. Rev. D, (1983), p. 679.

[39] S. M. BARNETT, R. P. CAMERON, AND A. M. YAO, *Duplex symmetry and its relation to the conservation of optical helicity*, Phys. Rev. A, 86 (2012), p. 013845.

[40] A. BARUT AND I. DURU, *Exact solutions of the dirac equation in spatially flat robertson-walker space-times*, Phys.Rev. D, (1987), p. 3705.

[41] M. BASTERO-GIL, A. BERERA, N. MAHAJAN, AND R. RANGARAJAN, *Power spectrum generated during inflation*, Phys. Rev. D, 87 (2013), p. 087302.

[42] J. D. BATES AND P. R. ANDERSON, *Effects of quantized scalar fields in cosmological spacetimes with big rip singularities*, Phys. Rev. D, 82 (2010), p. 024018.

[43] D. BAUMANN, *Tasi lectures on inflation*, arXiv:0907.5424v2, (2012).

[44] J. S. BELL AND R. JACKIW, *A pcac puzzle: π0■γγ in the σ-model*, Il Nuovo Cimento A (1965-1970), 60 (1969), pp. 47–61.

[45] M. A. BERGER, *Introduction to magnetic helicity*, Plasma Physics and Controlled Fusion, 41 (1999), p. B167.

[46] F. BEZRUKOV, D. GORBUNOV, AND M. SHAPOSHNIKOV, *On initial conditions for the hot big bang*, Journal of Cosmology and Astroparticle Physics, 2009 (2009), p. 029.

[47] N. BIRRELL, *The application of adiabatic regularization to calculations of cosmological interest*, Proy. R. Soc. B, (1978), p. 513.

[48] N. D. BIRRELL AND P. C. W. DAVIES, *Quantum Fields in Curved Space*, Cambridge University Press, Cambridge, UK, 1982.

[49] H. BONDI, M. VAN-DER BRUG, AND A. METZNER, *Gravitational waves in general relativity, vii. waves from axi-symmetric isolated system*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 269 (1962), pp. 21–52.

[50] C. BONVIN, C. CAPRINI, AND R. DURRER, *Magnetic fields from inflation: The transition to the radiation era*, Phys. Rev. D, 86 (2012), p. 023519.

[51] L. BREWIN, *Riemann normal coordinate expansions using cadabra*, Classical Quantum Gravity, (2009), p. 175017.

[52] T. S. BUNCH, *Bphz renormalization of $\lambda\phi^4$ field theory in curved spacetime*.

[53] T. S. BUNCH, *Adiabatic regularisation for scalar fields with arbitrary coupling to the scalar curvature*, J. Phys. A, (1980), p. 1297.

[54] T. S. BUNCH AND L. PARKER, *Feynman propagator in curved spacetime: A momentum-space representation*, Phys.Rev. D, (1979), p. 2499.

[55] M. G. CALKIN, *An invariance property of the free electromagnetic field*, American Journal of Physics, 33 (1965), pp. 958–960.

[56] R. CASADIO, F. FINELLI, M. LUZZI, AND G. VENTURI, *Higher order slow-roll predictions for inflation*, Phys. Lett., B625 (2005), pp. 1–6.

[57] C. CHEUNG, A. L. FITZPATRICK, J. KAPLAN, L. SENATORE, AND P. CREMINELLI, *The effective field theory of inflation*, Journal of High Energy Physics, 2008 (2008), p. 014.

[58] J. CHLUBA, J. HAMANN, AND S. P. PATIL, *Features and New Physical Scales in Primordial Observables: Theory and Observation*, Int. J. Mod. Phys., D24 (2015), p. 1530023.

[59] S. CHRISTENSEN, *Vacuum expectation value of the stress tensor in an arbitrary curved background: The covariant point-separation method*, Phys.Rev. D, (1976), p. 2490.

[60] S. M. CHRISTENSEN, *Regularization, renormalization, and covariant geodesic point separation*, Phys.Rev. D, (1978), p. 946.

[61] S. M. CHRISTENSEN AND M. J. DUFF, *Axial and Conformal Anomalies for Arbitrary Spin in Gravity and Supergravity*, Phys. Lett., B76 (1978), p. 571.

[62] S. M. CHRISTENSEN AND M. J. DUFF, *New Gravitational Index Theorems and Supertheorems*, Nucl. Phys., B154 (1979), pp. 301–342.

[63] A. DEL RIO, R. DURRER, AND S. P. PATIL, *Tensor bounds on the hidden universe*, To appear, (2018).

[64] A. DEL RIO, A. FERREIRO, J. NAVARRO-SALAS, AND F. TORRENTI, *Adiabatic regularization with a yukawa interaction*, Phys. Rev. D, 95 (2017), p. 105003.

[65] A. DEL RIO AND J. NAVARRO-SALAS, *Spacetime correlators of perturbations in slow-roll de sitter inflation*, Phys. Rev. D, 89 (2014), p. 084037.

[66] A. del Rio and J. Navarro-Salas, *Equivalence of adiabatic and dewitt-schwinger renormalization schemes*, Phys.Rev. D, (2015), p. 064031.

[67] A. del Rio and J. Navarro-Salas, *On the renormalization of ultraviolet divergences in the inflationary angular power spectrum*, Journal of Physics: Conference Series, 600 (2015), p. 012023.

[68] A. del Rio, J. Navarro-Salas, and F. Torrenti, *Renormalized stress-energy tensor for spin-1=2 fields in expanding universes*, Phys.Rev. D, (2014), p. 084017.

[69] S. Deser, *Off-shell electromagnetic duality invariance*, Journal of Physics A: Mathematical and General, 15 (1982), p. 1053.

[70] S. Deser and C. Teitelboim, *Duality transformations of abelian and non-abelian gauge fields*, Phys. Rev. D, 13 (1976), pp. 1592–1597.

[71] B. S. DeWitt, *Quantum field theory in curved spacetime*, Phys.Rep., (1975), p. 295.

[72] S. Dodelson, *Modern Cosmology*, Elsevier, 2003.

[73] A. Dolgov, I. Khriplovich, A. Vainshtein, and V. Zakharov, *Photonic chiral current and its anomaly in a gravitational field*, Nuclear Physics B, 315 (1989), pp. 138 – 152.

[74] J. S. Dowker, *Another discussion of the axial vector anomaly and the index theorem*, Journal of Physics A: Mathematical and General, 11 (1978), p. 347.

[75] J. S. Dowker and Y. P. Dowker, *Particles of arbitrary spin in curved spaces*, Proceedings of the Physical Society, 87 (1966), p. 65.

[76] M. J. Duff, *Twenty years of the weyl anomaly*, Classical and Quantum Gravity, 11, p. 1387.

[77] R. Durrer, *The Cosmic Microwave Background*, Cambridge University Press, 2008.

[78] R. Durrer, G. Marozzi, and M. Rinaldi, *Adiabatic renormalization of inflationary perturbations*, Phys. Rev. D, 80 (2009), p. 065024.

[79] G. Dvali and M. Redi, *Black hole bound on the number of species and quantum gravity at cern lhc*, Phys. Rev. D, 77 (2008), p. 045027.

[80] G. DVALI AND M. REDI, *Phenomenology of* 10^{32} *Dark Sectors*, Phys. Rev., D80 (2009), p. 055001.

[81] T. EGUCHI, P. B. GILKEY, AND A. J. HANSON, *Gravitation, gauge theories and differential geometry*, Physics Reports, 66 (1980), pp. 213–393.

[82] K. ENQVIST, T. MERINIEMI, AND S. NURMI, *Generation of the higgs condensate and its decay after inflation*, Journal of Cosmology and Astroparticle Physics, 2013 (2013), p. 057.

[83] K. ENQVIST, S. NURMI, S. RUSAK, AND D. J. WEIR, *Lattice calculation of the decay of primordial higgs condensate*, Journal of Cosmology and Astroparticle Physics, 2016 (2016), p. 057.

[84] A. FABBRI AND J. NAVARRO-SALAS, *Modeling Black Hole Evaporation*, Imperial College Press, 2004.

[85] A. FERREIRO AND J. NAVARRO-SALAS, *Pair creation in electric fields, anomalies, and renormalization of the electric current*, arXiv:1803.03247v1.

[86] D. G. FIGUEROA, *A gravitational wave background from the decay of the standard model higgs after inflation*, Journal of High Energy Physics, 2014 (2014), p. 145.

[87] D. G. FIGUEROA AND C. T. BYRNES, *The standard model higgs as the origin of the hot big bang*, Physics Letters B, 767 (2017), pp. 272 – 277.

[88] D. G. FIGUEROA, J. GARCIA-BELLIDO, AND F. TORRENTI, *Decay of the standard model higgs field after inflation*, Phys. Rev. D, 92 (2015), p. 083511.

[89] F. FINELLI, G. MAROZZI, G. P. VACCA, AND G. VENTURI, *Impact of ultraviolet regularization on the spectrum of curvature perturbations during inflation*, Phys. Rev. D, 76 (2007), p. 103528.

[90] L. H. FORD AND L. PARKER, *Infrared divergences in a class of robertson-walker universes*, Phys. Rev. D, (1977), p. 245.

[91] K. FUJIKAWA, *Path-integral measure for gauge-invariant fermion theories*, Phys. Rev. Lett., 42 (1979), pp. 1195–1198.

[92] K. FUJIKAWA, *Path integral for gauge theories with fermions*, Phys. Rev. D, 21 (1980), pp. 2848–2858.

[93] K. FUJIKAWA AND H. SUZUKI, *Path integrals and quantum anomalies*, 2004. Oxford University Press.

[94] S. A. FULLING, *Aspects of quantum field theory in curved spacetime*, Cambridge University Press, Cambridge, UK, 1989.

[95] S. A. FULLING AND L. PARKER, *Renormalization in the theory of a quantized scalar field interacting with a robertson-walker spacetime*, Ann. Phys. (N.Y.), (1974), p. 176.

[96] S. A. FULLING, L. PARKER, AND B. L. HU, *Conformal energy-momentum tensor in curved spacetime: Adiabatic regularization and renormalization*, Phys. Rev. D, (1974), p. 3905.

[97] J. GARCA-BELLIDO, D. G. FIGUEROA, AND J. RUBIO, *Preheating in the standard model with the higgs inflaton coupled to gravity*, Phys. Rev. D, 79 (2009), p. 063531.

[98] J. GARCIA-BELLIDO AND E. R. MORALES, *Particle production from symmetry breaking after inflation and leptogenesis*, Physics Letters B, 536 (2002), pp. 193 – 202.

[99] R. GEROCH, *Asymptotic Structure of Space-Time*, Springer US, Boston, MA, 1977, pp. 1–105.

[100] R. GEROCH, *Handbook of spacetime. chapter 15: Spinors*, Eds. A. Ashtekar and V. Petkov, Springer, (2015).

[101] S. GHOSH, *Creation of spin 1/2 particles and renormalization in flrw spacetime*, Phys.Rev. D, (2015), p. 124075.

[102] S. GHOSH, *Spin 1/2 field and regularization in a de sitter and radiation dominated universe*, Phys.Rev. D, (2016), p. 044032.

[103] P. B. GILKEY, *The spectral geometry of a riemannian manifold*, J. Differ. Geom., (1975), p. 601.

[104] G. F. GIUDICE, A. RIOTTO, I. TKACHEV, AND M. PELOSO, *Production of massive fermions at preheating and leptogenesis*, Journal of High Energy Physics, 1999 (1999), p. 014.

[105] I. S. GRADSHTEYN AND I. M. RYZHIK, *Table of Integrals, Series, and Products*, Academic Press, New York, 1994.

[106] P. B. GREENE AND L. KOFMAN, *Preheating of fermions*, Physics Letters B, 448 (1999), pp. 6 – 12.

[107] P. B. GREENE AND L. KOFMAN, *Theory of fermionic preheating*, Phys. Rev. D, 62 (2000), p. 123516.

[108] A. GUTH, *Inflationary universe: A possible solution to the horizon and flatness problems*, Phys. Rev., D (1981), p. 347.

[109] A. GUTH AND S.-Y. PI, *Fluctuations in the new inflationary universe.*, Phys. Rev. Lett., (1982), p. 1110.

[110] S. HABIB, C. MOLINA-PARIS, AND E. MOTTOLA, *Energy-momentum tensor of particles created in an expanding universe*, Phys. Rev. D, 61 (1999), p. 024010.

[111] S. W. HAWKING, *Particle creation by black holes*, Commun. Math. Phys., (1975), p. 199.

[112] S. W. HAWKING, *The development of irregularities in a single bubble inflationary universe.*, Phys. Lett., (1982), p. 295.

[113] S. W. HAWKING AND G. F. ELLIS, *The Large Scale Structure of the Universe*, Cambridge University Press, Cambridge, 1973.

[114] S. W. HAWKING, T. HERTOG, AND H. S. REALL, *Trace anomaly driven inflation*, Phys. Rev., D63 (2001), p. 083504.

[115] O. HEAVISIDE, *Xi. on the forces, stresses, and fluxes of energy in the electromagnetic field*, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 183 (1892), pp. 423–480.

[116] O. HEAVISIDE, *Electrical Papers*, Macmillan, 1893.

[117] B. HU AND L. PARKER, *Effect of gravitation creation in isotropically expanding universes*, Physics Letters A, 63 (1977), pp. 217 – 220.

[118] B. L. HU AND L. PARKER, *Anisotropy damping through quantum effects in the early universe*, Phys. Rev. D, 17 (1978), pp. 933–945.

[119] R. JACKIW, *Diverse topics in theoretical and mathematical physics*, 1995.

[120] J. D. JACKSON, *Classical electrodynamics*, Wiley, New York, NY, 3rd ed. ed., 1999.

[121] A. KAYA AND M. TARMAN, *Stress-energy tensor of adiabatic vacuum in friedmann-robertson-walker spacetimes*, J. Cosmol. Astropart. Phys., (2011), p. 040.

[122] T. KIMURA, *Divergence of axial-vector current in the gravitational field*, Progress of Theoretical Physics, 42 (1969), pp. 1191–1205.

[123] L. KOFMAN, A. LINDE, AND A. A. STAROBINSKY, *Reheating after inflation*, Phys. Rev. Lett., 73 (1994), pp. 3195–3198.

[124] L. KOFMAN, A. LINDE, AND A. A. STAROBINSKY, *Towards the theory of reheating after inflation*, Phys. Rev. D, 56 (1997), pp. 3258–3295.

[125] E. KOLB AND M. TURNER, *The early universe*, Addison-Wesley Publishing Company, USA, 1994.

[126] A. LANDETE, J. NAVARRO-SALAS, AND F. TORRENTI, *Adiabatic regularization for spin 1/2 fields*, Phys. Rev. D, (2013), p. 061501.

[127] A. LANDETE, J. NAVARRO-SALAS, AND F. TORRENTI, *Adiabatic regularization and particle creation for spin one-half fields*, Phys. Rev. D, (2014), p. 044030.

[128] A. R. LIDDLE AND D. H. LYTH, *Cosmological Inflation and Large-Scale Structure*, Cambridge University Press, 2006.

[129] A. D. LINDE, *A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems*, Phys. Lett., B (1982), p. 389.

[130] J. LOPEZ, *Asymptotic expansions of the whittaker functions for large order parameter.*, Meth. Appl. Anal., (1999), p. 249.

[131] J. MALDACENA, *Non-gaussian features of primordial fluctuations in single field inflationary models*, Journal of High Energy Physics, 2003 (2003).

[132] G. MAROZZI, M. RINALDI, AND R. DURRER, *On infrared and ultraviolet divergences of cosmological perturbations*, Phys. Rev. D, 83 (2011), p. 105017.

[133] J. MARTIN, C. RINGEVAL, AND V. VENNIN, *Encyclopaedia inflationaris*, Phys. Dark. Univ., 5 (2014), p. 75.

[134] J. MATYJASEK AND P. SADURSKI, *Stress-energy tensor of the quantized massive fields in friedman-robertson-walker spacetimes*, Phys. Rev. D, (2013), p. 104015.

[135] J. MATYJASEK, P. SADURSKI, AND M. TELECKA, *Vacuum polarization of the quantized massive fields in friedman-robertson-walker spacetime*, Phys. Rev. D, (2014), p. 084055.

[136] C. W. MISNER, K. S. THORNE, AND J. A. WHEELER, *Gravitation*, San Francisco: W. H. Freeman and Co., 1973.

[137] C. MOLINA-PARIS, P. R. ANDERSON, AND S. A. RAMSEY, *One-loop $\lambda\varphi^4$ field theory in robertson-walker spacetimes: Adiabatic regularization and analytic approximations*, Phys. Rev. D, 61 (2000), p. 127501.

[138] V. MUKHANOV, *Physical Foundations of Cosmology*, Cambridge University Press, 2005.

[139] V. F. MUKHANOV AND G. V. CHIBISOV, *Quantum fluctuations and a nonsingular universe*, JETP Lett., (1981), p. 532.

[140] M. NAKAHARA, *Geometry, topology and physics*, 2003. Bristol, UK: Hilger (1990) 505 p. (Graduate student series in physics).

[141] H. NIELSEN AND M. NINOMIYA, *The adler-bell-jackiw anomaly and weyl fermions in a crystal*, Phys. Lett. B, 130, (1983), 6.

[142] E. NOETHER, *Invariante variationsprobleme*, Nachr. v. d. Ges. d. Wiss. zu Gottingen, Math-phys. [English translation by M. A. Tavel. (http://arxiv.org/abs/physics/0503066v1)].

[143] P. O'DONNELL, *Introduction to 2-spinors in general relativity*, World Scientific, Singapore, 2003.

[144] T. PADMANABHAN, *Acceptable density perturbations from inflation due to quantum gravitational damping*, Phys. Rev. Lett., (1988), p. 2229.

[145] D. N. PAGE, *Particle emission rates from a black hole. ii. massless particles from a rotating hole*, Phys. Rev. D, 14 (1976), pp. 3260–3273.

[146] L. PARKER, *Amplitude of perturbations from inflation*, https://arxiv.org/pdf/hep-th/0702216.pdf.

[147] L. PARKER, *The creation of particles in an expanding universe*, Ph.D. thesis, Harvard University, 1966.

[148] L. PARKER, *Particle creation in expanding universes*, Phys. Rev. Lett., (1968), p. 562.

[149] L. PARKER, *Quantized fields and particle creation in expanding universes. 1*, Phys. Rev., (1969), p. 1057.

[150] L. PARKER, *Quantized fields and particle creation in expanding universes. 2*, Phys. Rev. D, (1971), p. 346.

[151] L. PARKER, *Particle creation and particle number in an expanding universe*, J. Phys. A, (2012), p. 374023.

[152] L. PARKER AND S. A. FULLING, *Adiabatic regularization of the energy-momentum tensor of a quantized field in homogeneous spaces*, Phys. Rev. D, (1974), p. 341.

[153] L. PARKER AND M. M. GLENZ, *Study of the spectrum of inflaton perturbations*, Phys. Rev. D, (2009), p. 063534.

[154] L. PARKER AND A. RAVEL, *Nonperturbative effects of vacuum energy on the recent expansion of the universe*, Phys. Rev. D, (1999), p. 063512.

[155] L. PARKER AND D. TOMS, *Explicit curvature dependence of coupling constants*, Phys. Rev. D, (1985), p. 2424.

[156] L. PARKER AND D. J. TOMS, *New form for the coincidence limit of the feynman propagator, or heat kernel, in curved spacetime*, Phys. Rev. D, (1985), p. 953.

[157] L. PARKER AND D. J. TOMS, *Quantum Field Theory in Curved Spacetime: Quantized Fields and Gravity*, Cambridge University Press, Cambridge, UK, 2009.

[158] L. PARKER AND D. VANZELLA, *Acceleration of the universe, vacuum metamorphosis, and the large-time asymptotic form of the heat kernel*, Phys. Rev. D, (2004), p. 104009.

[159] PAULANDERSON, *Effects of Quantum Fields on Singularities and Particle Horizons in the Early Universe. 4. Initially Empty Universes*, Phys. Rev., D33 (1986), p. 1567.

[160] M. PELOSO AND L. SORBO, *Preheating of massive fermions after inflation: analytical results*, Journal of High Energy Physics, 2000 (2000), p. 016.

[161] R. PENROSE, *Asymptotic properties of fields and space-times*, Phys. Rev. Lett., 10 (1963), pp. 66–68.

[162] R. PENROSE, *Gravitational collapse: The role of general relativity*, Riv. Nuovo Cim., 1 (1969), pp. 252–276. [Gen. Rel. Grav.34,1141(2002)].

[163] R. PENROSE, *The Road to Reality: a complete guide to the laws of the universe*, Jonathan Cape, London, 2004.

[164] R. PENROSE AND W. RINDLER, *Spinors and Space-Time*, vol. 1 of Cambridge Monographs on Mathematical Physics, Cambridge University Press, 1984.

[165] M. E. PESKIN AND D. V. SCHROEDER, *An Introduction to quantum field theory*, Addison-Wesley, 1995.

[166] PLANCK COLLABORATION, *Planck 2015 results - XI. CMB power spectra, likelihoods, and robustness of parameters*, AA, 594 (2016), p. A11.

[167] G. Y. RAINICH, *Electrodynamics in the general relativity theory*, Proceedings of the National Academy of Sciences, 10 (1924), pp. 124–127.

[168] R. RANGARAJAN, *Resolving the inflationary power spectrum*.

[169] M. REUTER, *Chiral anomaly of antisymmetric tensor fields*, Phys. Rev. D, 37 (1988), pp. 1456–1463.

[170] R.PENROSE, *Zero rest-mass fields including gravitation: asymptotic behaviour*, Proc. Roy. Soc. A, 284, (1965), 1397, pp. 159–203.

[171] B. T. S. AND P. C. W. DAVIES, *Quantum field theory in de sitter space: renormalization by point-splitting*, Proc. R. Soc. A, (1978), p. 117.

[172] R. K. SACHS, *Gravitational waves in general relativity viii. waves in asymptotically flat space-time*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 270 (1962), pp. 103–126.

[173] T. SAKAI, *On eigen-values of laplacian and curvature of riemannian manifold*, Tohoku Math. J., (1971), p. 589.

[174] M. D. SCHWARTZ, *Quantum Field Theory and the Standard Model*, Cambridge University Press, 2014.

[175] D. J. SCHWARZ, C. A. TERRERO-ESCALANTE, AND A. A. GARCIA, *Higher order corrections to primordial spectra from cosmological inflation*, Phys. Lett., B517 (2001), pp. 243–249.

[176] L. SENATORE AND M. ZALDARRIAGA, *On Loops in Inflation*, JHEP, 12 (2010), p. 008.

[177] M. S. SLOTH, *On the one loop corrections to inflation and the CMB anisotropies*, Nucl. Phys., B748 (2006), pp. 149–169.

[178] M. SREDNICKI, *Quantum Field Theory*, Cambridge University Press, 2007.

[179] A. A. STAROBINSKI AND P. S.-Y., *Fluctuations in the new inflationary universe.*, Phys. Lett., (1982), p. 175.

[180] A. A. STAROBINSKY, *Spectrum of relict gravitational radiation and the early state of the universe*, JETP Lett., 30 (1979), pp. 682–685.

[181] R. UTIYAMA AND B. S. DEWITT, *Renormalization of a classical gravitational field interacting with quantized matter fields*, Journal of Mathematical Physics, 3 (1962), pp. 608–618.

[182] A. VILENKIN AND L. H. FORD, *Gravitational effects upon cosmological phase transitions*, Phys. Rev. D, (1982), p. 1231.

[183] R. M. WALD, *Trace Anomaly of a Conformally Invariant Quantum Field in Curved Space-Time*, Phys. Rev., D17 (1978), pp. 1477–1484.

[184] R. M. WALD, *General relativity*, Chicago Univ. Press, Chicago, IL, 1984.

[185] R. M. WALD, *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*, University of Chicago Press, 1994.

[186] D.-G. WANG, Y. ZHANG, AND J.-W. CHEN, *Vacuum and gravitons of relic gravitational waves and the regularization of the spectrum and energy-momentum tensor*, Phys. Rev. D, 94 (2016), p. 044033.

[187] S. WEINBERG, *Feynman rules for any spin. ii. massless particles*, Phys. Rev., 134 (1964), pp. B882–B896.

[188] S. WEINBERG, *Quantum contributions to cosmological correlations*, Phys. Rev. D, 72 (2005), p. 043514.

[189] S. WEINBERG, *Cosmology*, Oxford University Press, Oxford, UK, 2008.

[190] S. WEINBERG, W. S, AND T. DE CAMPOS, *The Quantum Theory of Fields*, no. v. 1 in Quantum Theory of Fields, Vol. 2: Modern Applications, Cambridge University Press, 1995.

[191] C. M. WILL, *The confrontation between general relativity and experiment*, Living Reviews in Relativity, 17 (2014).

[192] R. P. WOODARD, *Perturbative quantum gravity comes of age*, International Journal of Modern Physics D, 23 (2014), p. 1430020.

[193] J. XIONG, S. K. KUSHWAHA, T. LIANG, J. W. KRIZAN, M. HIRSCHBERGER, W. WANG, R. J. CAVA, AND N. P. ONG, *Evidence for the chiral anomaly in the dirac semimetal na3bi*, Science, 350 (2015), pp. 413–416.