UNIVERSITY of CALIFORNIA
Santa Barbara

Addressing the Issues of Quantum Gravity in Generalized Quantum Mechanics

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Physics

by

John Thomas Whelan

Committee in charge:

Professor James B. Hartle, Chair
Professor Gary T. Horowitz
Professor Jeffrey D. Richman

December 1996

The dissertation of John Thomas Whelan is approved:

Chair

December 1996

Addressing the Issues of Quantum Gravity in Generalized Quantum Mechanics

In memory of my father,

Thomas Adrian Whelan.

1942–1993

Acknowledgements

This dissertation and the graduate career which it represents have been influenced in a number of ways by a number of people and institutions, and I hope I am forgiven if I neglect to thank any of them here.

For providing me with instruction and grounding in the field of physics, R. L. Sugar, J. Carlson, S. J. Peale, G. T. Horowitz, J. B. Hartle, J. D. Richman and J. L. Cardy. For providing the foundation on which my graduate career was built, Cornell University, the Astronomy Department, and especially S. W. Squyres.

For discussions on and off the topic of physics, my peers S. A. Rosenberg, J. Kikkawa, A. Chamblin, C. J. Fewster, and especially D. A. Craig. For inspiration from without, K. V. Kuchař, R. S. Tate, F. Dowker, D. Deutsch, H. D. Politzer, A. Anderson, L. Smolin, and B. Berger.

For making my year away from campus stimulating and enjoyable, the T-6 Group at Los Alamos National Laboratory, the Isaac Newton Institute for Mathematical Sciences, and the Department of Applied Mathematics and Theoretical Physics at Cambridge University.

For working out some of the calculations in Sec. III.5.4, and for years of advice, direction, inspiration, and encouragement, J. B. Hartle.

For untold emotional and practical support, Julie Bryce. For getting me where I am now and believing in me, my family, especially my mother Barbara Whelan.

Financially, this work was supported by NSF Grants PHY90-08502 and PHY95-07065, as well as an NSF Graduate Research Fellowship, and a University of California Regents Fellowship.

Curriculum Vitæ
John Thomas Whelan

April 24, 1969	Born, Kingston, NY
1989–1991	Undergraduate Researcher, Department of Astronomy, Cornell University, Ithaca, NY
1991	B.A. summa cum laude in Astronomy and with distinction in all subjects, Cornell University, Ithaca, NY
1991–1994	NSF Graduate Fellow, Department of Physics, University of California, Santa Barbara, CA
1994–1996	Graduate Student Researcher, Department of Physics, University of California, Santa Barbara, CA
1995	Teaching Assistant, Department of Physics, University of California, Santa Barbara, CA
1995–1996	Regents Special Fellow, University of California, Santa Barbara, CA

Publications

"Spacetime Alternatives in the Quantum Mechanics of a Relativistic Particle," Physical Review D, **50,** 6344-6356 (1994); gr-qc/9406029.

"Spacetime Coarse Grainings in Relativistic Particle Motion" in *The Proceedings of the Seventh Marcel-Grossman Conference*, edited by R. Ruffini and M. Keiser (World Scientific, Singapore 1995).

"Generalized Quantum Mechanics of Nonabelian Gauge Theories" Physical Review D, **53,** 3118-3146 (1996); hep-th/9507107

Abstract

Addressing the Issues of Quantum Gravity
in Generalized Quantum Mechanics

by

John Thomas Whelan

Hartle's generalized quantum mechanics (GQM) is applied to several toy models to gain insight into issues involved in the quantization of gravity, namely spacetime alternatives, gauge invariance, and the decoherence of spacetime.

Spacetime coarse grainings are studied in the GQM of a free relativistic particle. For a simple coarse graining and suitable initial conditions, tractable formulas are found for branch wave functions. Despite the nonlocality of the positive-definite version of the Klein-Gordon inner product, which means that nonoverlapping branches are not sufficient to imply decoherence, some initial conditions are found to give decoherence and allow the consistent assignment of probabilities.

The GQM of a nonabelian gauge theory is developed, and predictions made for certain alternatives, with particular attention given to those involving the constraint. In this way, the theory is compared to other quantum-mechanical descriptions of gauge theories in which the constraints are imposed *a priori*. The momentum space constraint is seen to vanish, both through a simple formal argument and a more careful description of the Lorentzian path integral as defined on a spacetime lattice. The configuration space realization of the constraint is shown to behave in a more complicated fashion. For some coarse grainings, I extend the known result from electromagnetism, that coarse grainings by values of the constraint either predict its vanishing or fail to decohere. However, sets of alternatives defined in terms of a more complicated quantity in electromagnetism are exhibited where definite predictions can be made disagreeing with the assumption that the constraints vanish. The configuration space sum-over-histories theory is exhibited in a manifestly Lorentz-invariant formulation.

The question of whether unobserved short-wavelength modes of the gravitational field can induce decoherence in the long-wavelength modes ("the decoherence of spacetime") is addressed using a simplified model of perturbative general relativity, in which the metric perturbation is replaced by a scalar field. For some long-wavelength coarse grainings, the Feynman-Vernon influence phase is found to be effective at suppressing the off-diagonal elements of the decoherence functional. The requirement that the short-wavelength modes be in a sufficiently high-temperature state places limits on the applicability of this perturbative approach.

Contents

List of Figures

Chapter I

Introduction

This dissertation is concerned with the implementation of one formalism, known as generalized quantum mechanics (GQM), designed to address some of the problems of quantum mechanics and its applications to gravity and cosmology. After a brief overview, in Sec I.1, of the issues motivating the research, this introductory chapter gives some useful background and develops the formalism in general terms before exploring the specific applications in Chapters II–IV. Secs. I.2 and I.3 address some of the interpretational issues in quantum theory which motivate generalized quantum mechanics. Sec. I.2 considers one problem which gives a simple mathematical motivation for the form of GQM, and then Sec. I.3 describes in more depth the problems with the standard Copenhagen interpretation of quantum mechanics and gives an explicit general formulation of generalized quantum mechanics. Sec. I.4 turns to the issue of the quantization of gravity, discussing various issues and techniques involved in attempts to quantize gravity, before describing, on a formal level, the approach used by GQM to handle the task. Finally, once the background and terminology have been developed, Sec. I.5 provides and outline for the rest of the dissertation.

Two appendices are included. Appendix I.A specifies some of the notation and conventions used in the dissertation. Appendix I.B reviews a few of the basics of General Relativity (GR), which successfully describes gravity on the classical level; particular attention is given to those aspects which are relevant to the rest of the dissertation.

I.1 Motivation

Each of the three chapters is primarily motivated by one issue in the quantization of gravity; two deal with the implementation of the quantization scheme, and the third is a physical question to be asked within that scheme.

I.1.1 Spacetime alternatives

One desirable property for a technique for quantizing gravity–the theory of spacetime–is that it treat time and space, equally. Thus traditional alternatives for which quantum mechanics predicts probabilities, such as whether a particle is in a given spatial region at a moment of time, ought to be generalized to include so-called spacetime alternatives, such as whether a particle enters a spacetime region at any point along its trajectory. Consideration of these spacetime alternatives in non-relativistic quantum mechanics (Sec. I.2) illustrates one of the features of generalized quantum mechanics, namely that probabilities cannot be predicted for all sets of alternatives, only for those which exhibit a lack of quantum mechanical interference between different alternatives.

Spacetime alternatives in non-relativistic quantum mechanics are considered in some depth in [1] and [2]; Chapter II examines spacetime alternatives in the quantum mechanics of a *relativistic particle*, which has the added complications that the path of the particle need not be single-valued in time, and which exhibits reparametrization invariance. In Chapter II, I exhibit a few simple sets of spacetime alternatives, some of which can be assigned probabilities, and calculate those probabilities.

I.1.2 Gauge invariance

General relativity is invariant under a set of transformations known as diffeomorphisms (Sec. I.B.3); while the relativistic particle exhibits a reparametrization invariance which thus makes it more similar to GR than the non-relativistic particle is, the invariance is not really central to the question of interest in Chapter II.

In Chapter III, however, gauge invariance is the principal focus in my study of the application of generalized quantum mechanics to non-Abelian gauge theories. (The detailed correspondence between the diffeomorphism invariance of GR and the reparametrization and gauge invariances of the theories considered in Chapters II and III is described in Sec. I.4.6.) Gauge invariant theories are conveniently described using more degrees of freedom than are needed to define the physical quantities of the theory, and many quantization methods use the constraints along with gauge-fixing conditions to isolate the "physical degrees of freedom". However, in a generalized quantum mechanics constructed via a sum over histories, it is possible to predict probabilities for different values of the constrained quantities, and that is what I do in several different contexts in Chapter III. Treating the constraints on the same footing as the classical equations of motion allows me to determine to what extent they come about naturally as a consequence of the quantum theory.

I.1.3 The decoherence of spacetime

Chapter IV addresses the following physical question: we know that measurements of the gravitational field on large scales produce (classical) predictions, so does generalized quantum mechanics allow us to make such predictions? In other words, do sets of alternatives describing only the long-wavelength modes of the gravitational field exhibit the non-interference between alternatives alluded to in Sec. I.1.1, known as decoherence? Decoherence is generally induced when the system of interest is coupled to some environment which carries away phase information, allowing the squares of probability amplitudes to sum linearly. In Chapter IV, I examine the question of whether sets of long-wavelength alternatives of the gravitational field can be made to decohere via their interaction with the unobserved short-wavelength modes.

I.2 Spacetime alternatives

There are a number of unsatisfying aspects of quantum mechanics which motivate the development of generalized quantum mechanics, and I will delve, in Sec. I.3, into the complex interpretational issues which lead to a desire to generalize quantum mechanics. But first, I'd like to consider a simple example of a sort of probability one might like to calculate in an extension of quantum mechanics, and see what modifications it naturally leads to.

Ordinary quantum mechanics can predict the value of an observable A at a time t; from a relativistic point of view, the surface of constant time t should be generalized to a spacelike surface σ. However, in a theory which includes quantum gravity, the metric itself will be behaving quantum mechanically, and the notion of spacelike separation of adjacent points will not be well-defined. Since we cannot tell if a surface is spacelike without evaluating (perhaps probabilistically) some of the variables of the theory, we should not restrict ourselves to the values of observables on spacelike surfaces, but rather attempt to define probabilities involving general regions of spacetime. This might be the average of a field over a spacetime region, or, as is considered in this secion, the probability that a single non-relativistic particle crosses, somewhere along its trajectory, a particular spacetime region S.

Since the sum-over-histories technique for calculating probabilities in quantum mechanics works, formally at least, with the entire trajectory of a system, it should be best suited to describing spacetime alternatives. Before proceding there, however, let us review the operator formalism of non-relativistic quantum mechanics.

I.2.1 Probabilities in quantum mechanics

In ordinary non-relativistic quantum mechanics [3], the outcome of a measurement is not generally predicted with certainty, but rather, the probabilities of possible outcomes can be predicted. Typically, the system is initially prepared in a state described by a wavefunction $\Psi(q', t')$, which can be written as a vector $|\Psi(t')\rangle$ in a Hilbert space with inner product

$$\langle\Phi|\Psi\rangle = \int dq\, \Phi^*(q)\Psi(q). \tag{I.2.1}$$

An observable A is described by a hermitian operator $\widehat{A}$ with eigenstates[1] $\{|a\rangle\}$ normalized so that $\langle a'|a\rangle = \delta(a - a')$. The probability that a measurement of A at a time t_i will fall into a range R is given by

$$p_{\mathrm{R}} = \int_R da\ |\langle a|\Psi(t_i)\rangle|^2\,, \tag{I.2.2}$$

[1] The case considered here is for an operator $\widehat{A}$ with a continuous spectrum, but the modifications for a discrete portion of the spectrum should be self-evident.

where the state $|\Psi(t_i)\rangle$ at time t_i is given by evolving the initial state[2] $|\Psi(t')\rangle$ using the operator $\widehat{H}$ corresponding to the Hamiltonian of the system:

$$|\Psi(t_i)\rangle = e^{-i(t_i-t')\widehat{H}}|\Psi(t')\rangle. \tag{I.2.3}$$

In particular, for the position operator $\widehat{Q}$ we recover the original interpretation of the wavefunction $\Psi(q,t_i) = \langle q|\Psi(t_i)\rangle$ as a probability amplitude.

It is also useful to write the probability (I.2.2) in terms of the projection onto the range R of eigenvalues of $\widehat{A}$

$$P_{\text{R}} = \int_R da|a\rangle\langle a|; \tag{I.2.4}$$

the projection of the initial state onto this range is called the branch wavefunction

$$|\Psi_{\text{R}}(t_i)\rangle = P_{\text{R}}|\Psi(t_i)\rangle, \tag{I.2.5}$$

which is a probability amplitude for A to lie in the range R:

$$p_{\text{R}} = \langle\Psi_{\text{R}}(t_i)|\Psi_{\text{R}}(t_i)\rangle. \tag{I.2.6}$$

Finally, this probability can be written in terms of the density matrix

$$\rho(t) = |\Psi(t)\rangle\langle\Psi(t)| \tag{I.2.7}$$

as

$$p_{\text{R}} = \text{Tr}[P_{\text{R}}\,\rho(t_i)] = \text{Tr}\left(e^{-i(t''-t_i)\widehat{H}}P_{\text{R}}e^{-i(t_i-t')\widehat{H}}\rho(t')e^{i(t_i-t')\widehat{H}}P_{\text{R}}e^{i(t''-t_i)\widehat{H}}\right). \tag{I.2.8}$$

The last form will ultimately lend itself most easily to generalization.

I.2.2 Sum-over-histories quantum mechanics

Another formulation of quantum mechanics, equivalent for many cases, was proposed by Feynman [4]. In the so-called sum-over-histories formulation, the transition amplitude between a position q' at a time t' and a position q'' at a time t'' is given by the sum over all paths $q(t)$ beginning at $q(t') = q'$ and ending at $q(t'') = q''$ of exp(i action):

$$\langle q''|e^{i(t''-t')\widehat{H}}|q'\rangle = \int_{q''q'} \mathcal{D}q e^{iS[q]}. \tag{I.2.9}$$

This *path integral* can be more explicitly described if the path is *skeletonized*, *i.e.*, the formal integral over continuous paths $q(t)$ is replaced with an approximation wherein the interval

[2] I will, unless otherwise noted, be working in the Schrödinger representation, in which the states, rather than the operators, evolve, as that is more convenient for the sum-over-histories formulation which I will eventually use. Even when states are written without time labels, they are to be thought of as Schrödinger states at the "default" times, *e.g.*, $|\Psi\rangle$ is shorthand for $|\Psi(t')\rangle$.

from t' to t'' is divided into $J+1$ sub-intervals of width $\delta t = \frac{t''-t'}{J+1} = \frac{T}{J+1}$, and the integral is reëxpressed in terms of $q^M = q(t' + M\delta t)$. In the limit $J \to \infty$, the skeletonized description should describe the integral over continuous paths.[3] Then the measure $\mathcal{D}q$ will be proportional to

$$\prod_{M=1}^{J} dq^M; \tag{I.2.10}$$

the proportionality constant, which depends upon J, is most easily calculated by starting with a canonical path integral and integrating out the momentum variables. The canonical path integral has the form

$$\langle q''|e^{i(t''-t')\widehat{H}}|q'\rangle = \int\limits_{q''q'} \mathcal{D}q\mathcal{D}p e^{iS_{\mathrm{can}}[q,p]} \tag{I.2.11}$$

in terms of the canonical action

$$S_{\mathrm{can}}[q,p] = \int_{t'}^{t''} dt[p\dot{q} - H(q,p)]. \tag{I.2.12}$$

It has the natural measure

$$\mathcal{D}q\mathcal{D}p = \left(\prod_{M=1}^{J} dq^M\right)\left(\prod_{M=0}^{J} \frac{dp^M}{2\pi}\right), \tag{I.2.13}$$

where the skeletonized momenta are given by $p^M = p\left(t' + [M+\frac{1}{2}]\delta t\right)$.

The probability rule (I.2.8) can be adapted to the sum-over-histories formulation by replacing the projection operator P_{R} with a restriction of the range of integration on the time slice corresponding to t_i. For example, if the operator $\widehat{A}$ is the position operator $\widehat{Q}$, the projection in (I.2.8) is replaced according to

$$\begin{aligned} e^{-i(t''-t)\widehat{H}} P_{\mathrm{R}} e^{-i(t-t')\widehat{H}} &= |q''\rangle \int\limits_{R} dq_i \int\limits_{q''q_i} \mathcal{D}q e^{i\int_{t_i}^{t''} dt L(t)} \int\limits_{q_i q'} \mathcal{D}q e^{i\int_{t'}^{t_i} dt L(t)} \langle q'| \\ &= |q''\rangle \int\limits_{\substack{q''q' \\ q_i\in R}} \mathcal{D}q e^{i\int_{t'}^{t''} dt L(t)} \langle q'| \end{aligned} \tag{I.2.14}$$

and so the branch wavefunction appearing in (I.2.6) can be written

$$\Psi_{\mathrm{R}}(q'',t'') = \int dq' \int\limits_{\substack{q''q' \\ q_i\in R}} \mathcal{D}q e^{iS[q]}\Psi(q',t'). \tag{I.2.15}$$

[3]Some of the stickier issues involved in making this correspondence will be discussed in Chapter III, especially Secs. III.3.3–III.3.5 and the discussion at the end of Sec. III.5.4. For now, however, all that is needed is the basic concept.

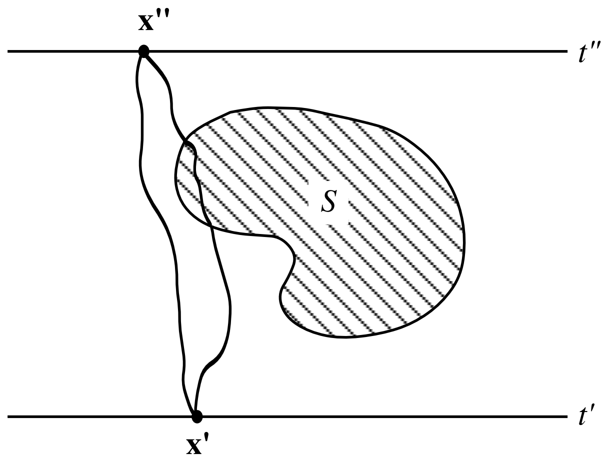


Figure I.1: Defining a set of spacetime alternatives in non-relativistic quantum mechanics. The path on the left never enters the spacetime region S and is thus in the class $c_{\overline{s}}$. The path on the right spends part of its trajectory in S and is thus in the class c_s.

Spacetime alternatives

Equation (I.2.15) describes the construction of a probability amplitude for a non-relativistic particle to pass through a given region of space R at a particular moment t_i of time by restricting the paths summed over to be those in which the particle does so. It is a short step to consider broadening the construction to produce the probability amplitude for the particle path to have any property by summing only over those paths which have that property. For example, as originally proposed by Feynman [4], one could define the branch wave function for the particle to pass through a space*time* region S at some point along its trajectory (see Fig. I.1) by restricting the sum to the class of paths c_s which pass through the region:

$$\Psi_s(q'', t'') = \int dq' \int_{q'' c_s q'} \mathcal{D}q e^{iS[q]} \Psi(q', t'). \tag{I.2.16}$$

The problem with this is that by the same reasoning, the probability amplitude that the path *never* enters S ought to be given by a sum over the class of paths $c_{\overline{s}}$ which never enter S:

$$\Psi_{\overline{s}}(q'', t'') = \int dq' \int_{q'' c_{\overline{s}} q'} \mathcal{D}q e^{iS[q]} \Psi(q', t'). \tag{I.2.17}$$

Since the sum over histories is linear, this branch wavefunction is obviously given by

$$\Psi_{\overline{s}}(q'', t'') = \Psi(q'', t'') - \Psi_s(q'', t'') \tag{I.2.18}$$

and hence

$$p_{\overline{s}} = \langle\Psi_{\overline{s}}|\Psi_{\overline{s}}\rangle = \langle\Psi|\Psi\rangle - 2\,\mathrm{Re}\langle\Psi|\Psi_s\rangle + \langle\Psi_s|\Psi_s\rangle = 1 - p_s - 2\,\mathrm{Re}\langle\Psi_{\overline{s}}|\Psi_s\rangle. \tag{I.2.19}$$

On the other hand, simple logic tells us that as the particle either passes through the region or doesn't, the sum of the two probabilities must be unity: $p_s + p_{\overline{s}} = 1$. This can only be true if

$$\mathrm{Re}\langle\Psi_{\overline{s}}|\Psi_s\rangle = 0. \tag{I.2.20}$$

That is always true in the case of a spatial region R at a time t_i by virtue of the construction (I.2.5), since $P_{\overline{\mathrm{R}}}P_{\mathrm{R}} = 0$, but for a spacetime region of the sort shown in Fig. I.1, will in general be true only for some values of the initial state $|\Psi\rangle$. Evidently (I.2.20) is a prerequisite for the squared norms of the amplitudes $|\Psi_s\rangle$ and $|\Psi_{\overline{s}}\rangle$ to behave as proper probabilities.

I.2.3 Generalized quantum mechanics

The simple demonstration in the previous section that spacetime probabilities cannot always be consistently assigned has provided the motivation for a more careful consideration of quantum mechanical probabilities and their applicability. I shall now describe the problem of spacetime alternatives in the language of one formalism designed to do this.

Sets of alternatives

The statement "the particle is in a spatial region R at time t_i", as well as the spacetime alternatives c_s and $c_{\overline{s}}$ considered in the previous section, are all examples of alternatives. An alternative is in general a statement which is either true or false for each *fine-grained history* of the system. In the current example of the relativistic particle, a fine-grained history is an arbitrary path $q(t)$; the sum over histories in (I.2.9) is over the set of all fine-grained histories. The class of fine-grained histories defined by an alternative is the set of all those histories for which the alternative is true. It is useful to consider a set of alternatives such that each history falls into exactly one of the classes (*i.e.*, the classes are mutually exclusive, and the set of classes is exhaustive). This partition of the histories is called a coarse graining; we write a given class or alternative as c_α and the complete set as $\{c_\alpha\}$.

Requirements on probabilities

One might try to define a probability p_α for each alternative c_α to occur. But we must require that the probability for a coarser-grained alternative which combines two disjoint alternatives ($c_\alpha \cap c_{\alpha'} = \emptyset$) be equal to the sum of the two individual probabilities:

$$p(c_\alpha \cup c_{\alpha'}) = p(c_\alpha) + p(c_{\alpha'}). \tag{I.2.21}$$

This means that if $\{c_\alpha\}$ is a complete set of alternatives for which probabilities $\{p_\alpha\}$ can be defined, they must sum to unity:

$$\sum_\alpha p_\alpha = 1. \tag{I.2.22a}$$

Also, if $\{c_{\overline{\alpha}}\}$ is a coarser-grained set formed by combining classes from the first set ($c_{\overline{\alpha}} = \bigcup_{\alpha\in\overline{\alpha}} c_\alpha$), one can assign probabilities to the alternatives in the second set by

$$p_{\overline{\alpha}} = \sum_{\alpha\in\overline{\alpha}} p_\alpha. \tag{I.2.22b}$$

These are known as the probability sum rules, and follow from (I.2.21).

Consistent assignment of probabilities

For any alternative c_α we can define a branch wave function by a sum restricted to the histories in the class:

$$\Psi_\alpha(q'',t'') = \int dq' \int_{q''\alpha q'} \mathcal{D}q e^{iS[q]}\Psi(q',t'); \tag{I.2.23}$$

however, the requirement (I.2.21) means that we can assign probabilities

$$p_\alpha = \langle\Psi_\alpha|\Psi_\alpha\rangle \tag{I.2.24}$$

to a complete set of alternatives if and only if each pair of alternatives in the set obeys

$$p_\alpha + p_{\alpha'} = (\langle\Psi_\alpha| + \langle\Psi_{\alpha'}|)(|\Psi_\alpha\rangle + |\Psi_{\alpha'}\rangle) = p_\alpha + p_{\alpha'} + 2\,\mathrm{Re}\langle\Psi_{\alpha'}|\Psi_\alpha\rangle \tag{I.2.25}$$

or

$$\mathrm{Re}\langle\Psi_{\alpha'}|\Psi_\alpha\rangle = 0. \tag{I.2.26}$$

This requirement (which is a condition on the alternatives and the states together) is called the condition of weak decoherence by Gell-Mann and Hartle [5], consistency by Griffiths [6] and Omnès [7] and non-interference by Yamada and Takagi [2].

Under Hartle's formulation of generalized quantum mechanics for the non-relativistic particle, the question of whether decoherence allows the assignment of probabilities to the alternatives in a coarse graining, and the values of those probabilities, are both encoded in the *decoherence functional*

$$D(\alpha,\alpha') = \langle\Psi_{\alpha'}|\Psi_\alpha\rangle. \tag{I.2.27}$$

If the off-diagonal elements of the decoherence functional [which is, in this description, a hermitian matrix with elements $\{D(\alpha,\alpha')\}$] vanish, then the diagonal elements can be identified as probabilities for the corresponding alternatives, $p_\alpha = D(\alpha,\alpha)$. If not, we cannot assign probabilities consistent with the sum rules (I.2.22). Mathematically, this is because the branch wavefunctions, which are probability amplitudes, add according to linear superposition, and the only way their squares–the probabilities–can add linearly as well is if the interference between two branches vanishes. I will return to the physical significance of this in Sec. I.3.

I.3 Formulations of quantum mechanics

Having illustrated the use of generalized quantum mechanics in the practical case of assigning probabilities to spacetime alternatives, I now consider the interpretational motivations involved in replacing the Copenhagen interpretation of quantum mechanics with a more general formalism. Then, in Sec. I.3.3, I describe GQM in its general form.

I.3.1 The two-slit experiment

The idea that not every set of alternatives can be assigned probabilities is one of the central interpretational challenges of quantum mechanics. The classic illustration is the two-slit experiment, wherein an electron gun is fired at an absorbing screen with two holes in it, beyond which lies another screen made up of detectors which will register the arrival of an electron. If the lower slit in the first screen is covered, the probability amplitude $\Psi_{\mathrm{U}}(x)$ for the electron to arrive at a position x on the screen can be calculated by propagating the initial wavefunction freely up to the screen, discarding the portion which lies outside the upper slit, then propagating the remaining wavefunction freely up to the final screen. The probability density for the electron to pass through the upper slit and arrive at a position x is thus $|\Psi_{\mathrm{U}}(x)|^2$. Similarly, if the upper slit is closed, we have a probability amplitude $\Psi_{\mathrm{L}}(x)$ and density $|\Psi_{\mathrm{L}}(x)|^2$ for the electron to pass through the lower slit and arrive at a point x on the second screen.

If both slits are opened, the principle of superposition tells us that the total probability amplitude for the electron to arrive at a position x on the second screen is $\Psi_{\mathrm{U}}(x)+\Psi_{\mathrm{L}}(x)$. This means that the total probability density

$$|\Psi_{\mathrm{U}}(x)+\Psi_{\mathrm{L}}(x)|^2 \neq |\Psi_{\mathrm{U}}(x)|^2+|\Psi_{\mathrm{L}}(x)|^2 \tag{I.3.1}$$

is not simply the sum of the probability densities for each slot alone, due to interference effects. This is true even if the electrons are released from the gun one at a time. Because of this interference, it is not possible to say that a given electron which lands on the second screen passed through one hole or the other. If a measurement is made at the first screen to detect the electron as it passes through one of the holes, the interference pattern is destroyed, and the probability density for arrival at point x *does* equal the expression on the right-hand side of (I.3.1). This means that something more complicated than simple Schrödinger evolution of the wavefunction must occur when it interacts with the measuring apparatus. All of these phenomena are confirmed by experiment.

In the language of Sec. I.2.3, this means that we cannot assign probabilities to a set of alternatives which includes

- passage through slit not observed/electron passes through upper slit/electron arrives at position $x \pm \Delta x$
- passage through slit not observed/electron passes through lower slit/electron arrives at position $x \pm \Delta x$,

while we can assign probabilities to the set

- passage through slit not observed/electron arrives at position $x \pm \Delta x$
- passage through slit observed/electron passes through upper slit/electron arrives at position $x \pm \Delta x$
- passage through slit observed/electron passes through lower slit/electron arrives at position $x \pm \Delta x$.

I.3.2 The Copenhagen interpretation

The standard method for resolving this issue, and those like it, was developed primarily by Niels Bohr at the Copenhagen Institute (see Chapter 2, Sec. 6-8 of [3] for more details), and is known as the Copenhagen interpretation. In the Copenhagen interpretation, the wavefunction for a quantum system evolves freely according to (I.2.3) until a measurement is made, say at time t_i. The measurement not only yields an answer with probabilities given by (I.2.5–I.2.6), it also alters the wavefunction. If the alternative corresponding to P_{R} is the result of the measurement, the wavefunction becomes [*cf.* (I.2.5)]

$$\frac{|\Psi_{\mathrm{R}}(t_i)\rangle}{\sqrt{\langle\Psi_{\mathrm{R}}(t_i)|\Psi_{\mathrm{R}}(t_i)\rangle}} = \frac{P_{\mathrm{R}}|\Psi(t_i)\rangle}{\sqrt{\langle\Psi(t_i)|P_{\mathrm{R}}|\Psi(t_i)\rangle}}. \tag{I.3.2}$$

This is known as "collapse of the wavefunction" or "reduction of the wavepacket".

The Copenhagen interpretation explains the results of the two-slit experiment nicely; if no observation of the wavefunction is made at the first screen, it propagates freely through both slits, and the arrival pattern on the second screen exhibits quantum-mechanical interference effects, while if the electron is observed to go through one slit, the wavefunction is collapsed by a projection onto that slit. Even if the result of the first measurement is ignored, we cannot recover the quantum-mechanical interference pattern; the probabilities for different ranges of arrival position x are given by the rule (I.2.8), where the density martix involved is now the mixed state

$$\rho = |\Psi_{\mathrm{U}}\rangle p_{\mathrm{U}}\langle\Psi_{\mathrm{U}}| + |\Psi_{\mathrm{L}}\rangle p_{\mathrm{L}}\langle\Psi_{\mathrm{L}}|, \tag{I.3.3}$$

where $p_{\mathrm{U}} = \langle\Psi_{\mathrm{U}}|\Psi_{\mathrm{U}}\rangle$ and $p_{\mathrm{L}} = \langle\Psi_{\mathrm{L}}|\Psi_{\mathrm{L}}\rangle$ are the probabilities for the outcome of the ignored measurement.

The Copenhagen interpretation is quite successful in predicting and explaining the results of laboratory experiments, for which it was designed. It cannot, however, provide a complete quantum-mechanical description of an arbitrary system. Setting aside the aesthetic objection that the two evolution laws (I.2.3) and (I.3.2) are of markedly different nature, a major flaw is the fuzzy notion of a measurement. What are the criteria for an object with which the quantum system interacts to qualify as an observer? There are many metaphysical answers to that question, but the most conservative one would seem to be that the "measuring apparatus" is a classical object with many internal degrees of freedom. One possible outlook is that objects in the classical world are described by classical physics, quantum-mechanical objects by quantum physics, and the reduction of the wavefunction describes what happens when the two regimes interact. However, results such as Ehrenfest's theorem, which show that the laws of classical physics are the limits of the laws of quantum physics when the actions involve become much larger than $\hbar$, lead us to believe that both the quantum and classical worlds should be described on the fundamental level by the same set of physical laws. From that perspective, it is natural to think of the Copenhagen interpretation as an approximation, wherein the complex interaction between the quantum system and the classical object's many degrees of freedom is described simply by the collapse of the wavefunction. (The details of this question touch on the rich subject of measurement theory, and the reader is again directed to [3] for a more in depth discussion.)

At any rate, a more general formulation of quantum mechanics is required to predict probabilities for closed systems without an external classical observer. This could either be a description of an open system together with an external observer, where both the "system" and the "classical observer" are combined to produce a single closed system, treated quantum mechanically, or a system like the early universe, where there is no classical subsystem to perform the measurements.

I.3.3 The generalized quantum mechanics formalism

One prescription for replacing the notions of "observers" and "measurements" is generalized quantum mechanics, as described in Sec. I.2.3 for a system consisting of a single non-relativistic particle. Probabilities can be assigned to any set of alternatives which decohere, which is a mathematical property of the system; an external classical observer is not required. The formulation given here was codified by Hartle [5], expanding upon earlier work of Gell-Mann and Hartle [8], inspired by the work of Żurek [9] on decoherence and related to the previous independent work of Griffiths [6] and Omnès [7].

Fundamental elements

In its most general form, a generalized quantum mechanics requires three things, which are generalizations of the objects described in Sec. I.2.3:

1. a specification of the fine-grained histories of the system,
2. a rule defining how one is allowed to combine those fine-grained histories into coarse-grained classes, and
3. a decoherence functional, which describes both the interference between pairs of those classes, as well as probabilities of individual classes when the interference terms vanish.

In general, the decoherence functional, whose elements $D(\alpha, \alpha')$ can be thought of as elements of a complex matrix for each coarse graining, can be constructed in an arbitrary fashion. It must, however, obey the following rules:

- "Hermiticity":
$$D(\alpha', \alpha) = D(\alpha, \alpha')^*; \tag{I.3.4a}$$
- positivity of diagonal elements:
$$D(\alpha, \alpha) \geq 0; \tag{I.3.4b}$$

- normalization:

$$\sum_{\alpha}\sum_{\alpha'} D(\alpha,\alpha') = 1; \tag{I.3.4c}$$

- superposition: If $\{c_{\overline{\alpha}}\}$ is a coarse graining constructed by combining classes in $\{c_\alpha\}$ to form larger classes ("coarser graining"), i.e., $c_{\overline{\alpha}} = \bigcup_{\alpha\in\overline{\alpha}} c_\alpha$, the decoherence functional for $\{c_{\overline{\alpha}}\}$ can be constructed from the one for $\{c_\alpha\}$ by

$$D(\overline{\alpha},\overline{\alpha}') = \sum_{\alpha\in\overline{\alpha}}\sum_{\alpha'\in\overline{\alpha}'} D(\alpha,\alpha'). \tag{I.3.4d}$$

The diagonal elements $D(\alpha,\alpha)$ can be interpreted as probabilities if the coarse graining $\{c_\alpha\}$ obeys some decoherence condition. The minimal one neccessary to ensure the probability sum rules (I.2.22) is *weak* decoherence, *i.e.,* vanishing of off-diagonal elements of $\operatorname{Re} D(\alpha,\alpha')$. In practice, it is useful to impose the decoherence condition only approximately:[4]

$$\operatorname{Re} D(\alpha,\alpha') \approx \delta_{\alpha\alpha'} p_\alpha, \tag{I.3.5}$$

with the diagonal elements then being interpreted as approximate probabilities whose accuracy is limited by the size of the off-diagonal elements. Another decoherence condition, more useful from a mathematical point of view, is the *medium* decoherence condition

$$D(\alpha,\alpha') \approx \delta_{\alpha\alpha'} p_\alpha. \tag{I.3.6}$$

I will adopt the practice of using "decoherence" with no modifier to mean medium decoherence. (There are stronger definitions of decoherence [11], but they will not be needed here.)

The notion of decoherence replaces the Copenhagen idea of a measurement in determining when classical probabilities can be assigned. In the case of a classical measuring apparatus interacting with a quantum system, it had been shown [12, 9] that the classical object can carry away phase information and induce decoherence between alternatives of the quantum system.

Operator generalized quantum mechanics

As an example of how a decoherence functional is constructed, return to the example of nonrelativistic operator quantum mechanics discussed in Sec. I.2.1. The formula for the decoherence functional should reduce to the probability formula (I.2.8) for the case of projections on a single time slice. A more general coarse graining can consist of a series of n time slices $\{t_i\}$, each of which has a complete set of projection operators $P^i_{\alpha_i}$ obeying

$$\sum_{\alpha_i} P^i_{\alpha_i} = \mathbb{I} \tag{I.3.7a}$$

$$P^i_{\alpha_i} P^i_{\alpha'_i} = \delta_{\alpha_i\alpha'_i} P^i_{\alpha_i}. \tag{I.3.7b}$$

[4] Readers disquieted by this concept will be reassured by the result [10] that for every approximately decohering coarse graining, there is a nearby coarse graining which decoheres exactly.

(The superscript i labels the set of projection operators, while the subscript α_i labels the particular projection within the set.) A class c_α is now labelled by the sequence of values $\{\alpha_1, \ldots, \alpha_n\}$, which we abbreviate in the shorthand α, and the place of a projection operator for c_α is taken by the chain of projections

$$C_\alpha = e^{-i(t''-t_n)\widehat{H}} P^n_{\alpha_n} e^{-i(t_n-t_{n-1})\widehat{H}} P^{n-1}_{\alpha_{n-1}} \ldots P^2_{\alpha_2} e^{-i(t_2-t_1)\widehat{H}} P^1_{\alpha_1} e^{-i(t_1-t')\widehat{H}}. \tag{I.3.8}$$

This is used in the formula for the decoherence functional,

$$D(\alpha, \alpha') = \mathrm{Tr}[C_\alpha \rho C^\dagger_{\alpha'}]. \tag{I.3.9}$$

It is straightforward to show that when the class operator consists of a single projection ($C_\alpha = e^{-i(t''-t_i)\widehat{H}} P_\alpha e^{-i(t_i-t')\widehat{H}}$), the decoherence functional reduces exactly to the diagonal form in (I.3.6) with the probabilities given by (I.2.8).

One final generalization can be made, by replacing the initial condition $\rho(t')$ with initial and final states described by $\rho' = \rho'(t')$ and $\rho'' = \rho''(t'')$, respectively, to give

$$D(\alpha, \alpha') = \frac{\mathrm{Tr}[\rho'' C_\alpha \rho' C^\dagger_{\alpha'}]}{\mathrm{Tr}[\rho'' e^{-i(t''-t')\widehat{H}} \rho' e^{i(t''-t')\widehat{H}}]}. \tag{I.3.10}$$

This is construction is most useful when considering the origins of the "arrow of time" [13], since it provides a time-reversal invariant formulation of the theory. The form (I.3.10) is recovered as a consequence of the final condition $\rho'' = \mathbb{I}$, known as the condition of future indifference.

As a consequence of (I.3.7) the class operator obeys

$$C_u = \sum_\alpha C_\alpha = e^{-i(t''-t')\widehat{H}}. \tag{I.3.11}$$

In the sum-over-histories formulation, the chain of projections C_α is replaced by a *class operator* whose matrix elements are defined via a path integral, as described below.

Sum-over-histories generalized quantum mechanics

In a sum-over-histories formalism, one replaces objects defined by operators on a Hilbert space with corrseponding quantities defined by functions (or functionals) of the configuration space coördinates of the system, which are written generically as q. Along those lines, the initial density matrix ρ' is replaced by a set of wave function(al)s[5] $\{\Psi_j(q')\}$ with corresponding non-negative weights (or "probabilities") $\{p'_j\}$. (In a Hilbert space theory this would mean defining $\rho' = \sum_j |\Psi_j\rangle p'_j \langle\Psi_j|$.) Similarly, the final state is now defined by a set of wave function(al)s $\{\Phi_i(q'')\}$ and weights $\{p''_i\}$, which replace the density matrix ρ''.

With these conventions, the definition (I.3.10) for the decoherence functional is replaced by

$$D(\alpha, \alpha') = \frac{\sum_{i,j} p''_i \langle\Phi_i|C_\alpha|\Psi_j\rangle\langle\Phi_i|C_{\alpha'}|\Psi_j\rangle^* p'_j}{\sum_{i,j} p''_i \left|\langle\Phi_i|C_u|\Psi_j\rangle\right|^2 p'_j}. \tag{I.3.12}$$

[5] I have suppressed the labels t' and t'' here, in part because I wish to describe more general definitions of the endpoints of histories. For example, the initial state might be attached on a surface σ' in spacetime.

Here the quantity $\langle\Phi_i|C_\alpha|\Psi_j\rangle$ is analogous to a matrix element of the class operator for the class c_α, but it is constructed by a sum over the histories in the class c_α, with the initial and final wave function(al)s Ψ_j and Φ_i attached at the endpoints of the history (which are included in the sum). Schematically:

$$\langle\Phi_i|C_\alpha|\Psi_j\rangle = \Phi_i(q'') \circ \langle q''\|C_\alpha\|q'\rangle \circ \Psi_j(q') \tag{I.3.13a}$$

$$\langle q''\|C_\alpha\|q'\rangle = \sum_{\substack{\text{history}\in\alpha \\ \text{endpts}=q''q'}} e^{iS[\text{history}]}. \tag{I.3.13b}$$

The inner product $\circ$ (which is in general necessary to ensure a non-divergent construction for cosntrained theories) used to attach the initial and final wave function(al)s must be Hermitian[6] but not necessarily positive definite.

This construction satisfies the requirements (I.3.4) for a decoherence functional with positivity of diagonal elements (I.3.4b) holding as long as the weights $\{p'_j\}$ and $\{p''_i\}$ are non-negative. Note that the inner product $\circ$ need not be positive definite to ensure positivity of the decoherence functional. The superposition property (I.3.4d) holds because the class operators are constructed linearly, and thus satisfy their own superposition property:

$$\langle q''\|C_{\overline{\alpha}}\|q'\rangle = \sum_{\alpha\in\overline{\alpha}} \langle q''\|C_\alpha\|q'\rangle \tag{I.3.14a}$$

$$\sum_\alpha \langle q''\|C_\alpha\|q'\rangle = \langle q''\|C_u\|q'\rangle. \tag{I.3.14b}$$

It is convenient to refer to a "class operator C_α" even in the sum-over-histories theory, and when I do, I mean the object defined by (I.3.13). C_u is the class operator corresponding to the class c_u of all paths, which is just the propagator.

Another way of expressing (I.3.12–I.3.13) is to write down the decoherence functional between two classes which are ultimately fine-grained, so that each class contains a single history $q(t)$. This is then truly a *functional* of its two arguments, given by

$$D[q_1,q_2] = \frac{\rho''(q''_2,q''_1)e^{i(S[q_1]-S[q_2])}\rho'(q'_1,q'_2)}{\int\mathcal{D}q_1\int\mathcal{D}q_2\rho''(q''_2,q''_1)e^{i(S[q_1]-S[q_2])}\rho'(q'_1,q'_2)}, \tag{I.3.15}$$

where the function(al)s corresponding to the initial and final density matrices are just

$$\rho'(q'_1,q'_2) = \sum_j \Psi^*_j(q'_1)p'_j\Psi_j(q'_2) \tag{I.3.16a}$$

$$\rho''(q''_1,q''_2) = \sum_i \Phi^*_i(q''_1)p''_i\Phi_i(q''_2) \tag{I.3.16b}$$

The coarse-grained decoherence functional is then given by (I.3.4d) as

$$D(\alpha_1,\alpha_2) = \int\limits_{\alpha_1}\mathcal{D}q_1\int\limits_{\alpha_2}\mathcal{D}q_2 D[q_1,q_2]. \tag{I.3.17}$$

[6]By which I mean $\Phi\circ\Psi = (\Psi\circ\Phi)^*$.

Field theory In the previous section, I described formally the sum-over-histories formulation of generalized quantum mechanics where the configuration space variables of the theory were either some set of coördinates $\{q_i(t)\}$ or a field $q(\mathbf{x},t)$ on the background spacetime. One can of course always switch between the two by imagining the (D-dimensional) space spanned by $\mathbf{x}$ to be divided up into a small lattice with lattice spacing δx and using expressions like

$$\delta^D(\mathbf{x}-\mathbf{x}') \Leftrightarrow \frac{\delta_{\mathbf{x}\mathbf{x}'}}{(\delta x)^D} \tag{I.3.18}$$

to make the correspondence

$$q_{\mathbf{x}}(t) \Leftrightarrow (\delta x)^{D/2} q(\mathbf{x},t), \tag{I.3.19}$$

but it is useful to be explicit about the form of some of the field theory expressions.

The initial condition $q(t')=q'$, formerly imposed at the initial time t', is now imposed on the initial surface labelled by t' as $q(\mathbf{x},t')=q'(\mathbf{x})$. The initial wave functional[7] $\Psi[q';t')$ is a functional of the function $q'(\mathbf{x})$. Similarly, the final wave functional $\Phi[q'';t'')$ attached to $\langle\Phi|C_\alpha|\Psi\rangle$ is a functional of $q''(\mathbf{x})=q(\mathbf{x},t'')$.

I.4 Quantum gravity

I.4.1 Motivation

We have now come to the main purpose of this dissertation: applying the lessons learned about quantum theory in Secs. I.2–I.3 to the quantization of the geometric theory of gravity, reviewed in Appendix I.B. Two obvious questions arise as to why a quantum theory of gravity is necessary. Firstly, the classical theory of gravity given by GR has held up under every experimental test required of it. Perhaps the classical theory really is the true fundamental theory.

The second question is a practical one. One might ask if a quantum theory of gravity is relevant to any observable phenomena. On dimensional grounds, we expect quantum effects to become important at length scales around the Planck length

$$\ell_p = G^{1/2} = \sqrt{\frac{G\hbar}{c^3}}, \tag{I.4.1}$$

which we see from the form with c and $\hbar$ inserted explicitly is the unique combination of G, c and $\hbar$ with units of length. Since G and $\hbar$ are so small in practical units, and c so large, ℓ_p turns out to have the incredibly small value of

$$\ell_p = 1.6\times 10^{-33}\,\mathrm{cm}, \tag{I.4.2}$$

Which is far below the resolution limits of any experimental device. The corresponding energy $E_p = 1.3\times 10^{19}\,\mathrm{GeV}$ is beyond the range of any conventional particle accelerator that could be built on the surface of the earth.

[7] The mismatched parentheses indicate that Ψ is a functional of $q'(\mathbf{x})$ and a function of t'; see Appendix I.A.

Addressing the first question first, there are a number of reasons to believe that there is a fundamental quantum theory of gravity underlying classical GR [14]. First of all, the other forces of nature can be explained on a fundamentally quantum level (electroweak theory and quantum chromodynamics), and it seems odd for one force to be fundamentally classical while the others are quantum mechanical. This is especially true if one has aspirations of unifying gravity with the other forces as Maxwell's classical electromagnetism unified electricity with magnetism, the Weinberg-Salam electroweak theory [15] unified quantum electrodynamics with the weak interaction, and a hypothetical grand unified theory might unify electroweak theory with quantum chromodynamics. To put an even finer point on it, the Einstein equation (I.B.22) couples the curvature tensor $G_{\mu\nu} = R_{\mu\nu} - \frac{g_{\mu\nu}}{2}R$, which is supposed to behave classically, with the stress-energy tensor $\{T_{\mu\nu}\}$, which is generated by quantum-mechanical matter fields. The straightforward fix to this problem, replacing the quantum stress-energy tensor with its expectation value in the state $|\Psi\rangle$ to give

$$G_{\mu\nu} = 8\pi G\langle\Psi|T_{\mu\nu}|\Psi\rangle, \tag{I.4.3}$$

produces useful results in the field of semiclassical gravitational physics, but still suffers conceptual problems when $\{T_{\mu\nu}\}$ describes a superposition of widely differing matter configurations [14].

Another reason why classical GR ought not to be taken as a fundamental theory is its prediction of spacetime singularities [16]. There are theorems which demonstrate that generic non-singular initial data can collapse into a black hole, which contains a spacetime singularity at which the curvature diverges; in the presence of this divergence, GR is unable to predict the subsequent evolution of the system. There is an escape from the problem of singularities formed by collapse, in that most if not all such singularities predicted by classical GR will be surrounded by "event horizons", regions from which no timelike or lightlike signal can escape to the external spacetime. This means that it doesn't matter that we don't know how to interpret the future evolution of such singularities, as they will not lie in the past light cone of any exterior observer. However, the same singularity theorems predict that evolving our expanding universe backwards in time will always lead to an initial singularity, or "big bang". This singularity is on everyone's light cone, and thus General Relativity predicts that our universe originated in a phenomenon it is unable to describe adequately.

Turning now to the question of whether a fundamental quantum theory of gravity, if it exists, is of practical interest, the singularity theorems again argue for the affirmative. As the system evolves towards an infinite-curvature singularity, the curvature will go through arbitrarily large values, and when it exceeds the Planck curvature ℓ_p^{-2}, quantum gravitational effects should become relevant. Again, GR is probably safe from ordinary astrophysical black holes, as the curvature at the event horizon of a black hole is of the order of $(GM_{\mathrm{BH}})^{-2} = m_p^4/M_{\mathrm{BH}}^2$ [17], so that it will only reach the Planck curvature in observable regions if $M_{\mathrm{BH}} \lesssim m_p$. Although semiclassical physics predicts that black holes will lose mass with time due to Hawking evaporation, the time required for a solar mass black hole to evaporate to Planck size will be [14] on the order of $\left(\frac{M_\odot}{m_p}\right)^3 t_p \sim 1.3 \times 10^{63}\,\mathrm{yr}$, fifty-four orders of magnitude greater than the age of the universe. However, the initial singularity predicted by comsmology again compels us to consider quantum gravity, as the early universe must have had curvatures in the Planck regime, and hence the "initial condition" from which the classical evolution of our universe began came about as a result of quantum gravitational phenomena. An understanding of these phenomena could allow us to explain, for example, the recently-observered inhomogeneities in the cosmic microwave background radiation (CMBR) [18].

In addition, there are the obvious intellectual motivations to pursue a fully quantum-mechanical theory of gravity. If there is a more fundamental theory of gravity than classical GR, it is the job of theoretical physics to formulate it. While this is difficult in the absence of experimental data[8] to distinguish between competing theories, the inherent challenges in formulating a consistent theory in the first place (see the following section) provide their own tests for what is a "good" theory of gravity. And finally, just as special and general relativity, as well as quantum mechanics, gave physicists new insight into the nature of the universe, the pursuit of a quantum theory of gravity will lead to new ways of thinking about the universe. String theory and the loop representation (see Sec. I.4.4) certainly fall into that category, as does generalized quantum mechanics itself!

I.4.2 Difficulities in quantizing gravity

The dual role of geometry

The first obstacle to be overcome in formulating a quantum theory of gravity is the fundamental role played by the spacetime geometry. Replacing the external gravitational field of Newton with the beautiful geometric theory summarized in Appendix I.B, prevents a straightforward application of the usual recipies for quantization. For example, [14] the causality condition

$$[\widehat{g}_{\mu\nu}(x), \widehat{g}_{\lambda\sigma}(x')] = 0, \quad x \text{ and } x' \text{ spacelike separated} \tag{I.4.4}$$

is difficult to interpret when the geometry itself is behaving quantum mechanically. If the spacetime metric is now an operator $\{\widehat{g}_{\mu\nu}\}$ rather than a classical quantity with a well-defined value, what metric is being used to determine whether or not x and x' are spacelike separated? Clearly, any fully quantum-mechanical theory of gravity must be constructed with some deeper insight as to how to quantize geometry.

The problem of time

A related issue is the "problem of time", namely that the time direction enjoys a special role in most methods of quantization, inconsistent with the Lorentz invariance of special relativity and the diffeomorphism invariance of general relativity. States of the system are defined on constant-time surfaces, as are spaces of eigenstates onto which projections defining measurements are made. Commutation relations and canonical methods are defined using conjugate momenta, which also single out a preferred time direction. Even when surfaces of constant time are generalized in these definitions to include arbitrary timelike surfaces, there is still a problem, as the question of whether two spacetime points are timelike or spacelike separated will be ill-defined if the metric is behaving fully quantum mechanically.

It is these sorts of considerations which make the spacetime alternatives of Sec. I.2 so appealing, and they will be the focus of Chapter II.

[8] Other, that is, than the indirect information afforded us by cosmological observations.

Renormalization

Another problem which faces would-be quantum theories of gravity is that of renormalizibility. Interacting and self-interacting field theories are only fully understood at the perturbative level, where quantities of interest are expressed as perturbation series in the coupling contant. Unfortunately, terms in the perturbation series beyond zero order will generally be equal to divergent integrals over the momenta of intermediate particles. If some sort of artificial cutoff is introduced to remove this divergence, the theory will then depend on these cutoff parameters (and the values of physical constants will diverge as those cutoffs are taken to infinity). The dependence on the cutoffs can be cancelled out if additional terms ("counter-terms") are added to the action which contain the same divergences. In general, then, the action will then have a number of arbitrary parameters on which the theory will depend.

The one way out of this jam is if suitable counter-terms can be constructed which have the same form as terms already in the action. Then the new parameters can be absorbed into redefinitions of the old ones. This scheme only works if the coupling constant is dimensionless, so that the divergent parts of simple diagrams can cancel those of more complex ones including additional verticies. If the coupling constant has dimensions which are some inverse power of length (positive mass dimensions), later terms in the expansion will involve more factors of momentum in the denominator, and won't diverge in the ultraviolet limit. However, if the coupling constant has positive length (negative mass dimensions) the theory should be non-renormalizable.

A simple way to analyze the expected behavior of GR under perturbative quantization is to expand the metric about some classical solution, replacing $g_{\mu\nu}$ by $g_{\mu\nu}+\gamma_{\mu\nu}$. The gravitational action in this case is calculated in Appendix IV.A to chapter IV, and is given by [*cf.* (IV.A.41)]

$$S=\frac{1}{16\pi G}\int\sqrt{|g|}\,d^4x\left[-\frac{1}{4}(\nabla_\lambda\gamma_{\mu\nu})(\nabla^\lambda\gamma^{\mu\nu})+\frac{1}{2}\gamma^{\mu\nu}R_{\mu\lambda\nu\sigma}\gamma^{\lambda\sigma}+\mathcal{O}(\{\gamma_{\mu\nu}\}^3)\right]. \tag{I.4.5}$$

In order that the lowest order term resemble that of a free field [*cf.* (I.B.15)] we need to quantize the field $\frac{1}{\sqrt{16\pi G}}\gamma_{\mu\nu}=\frac{1}{\ell_p 4\sqrt{\pi}}\gamma_{\mu\nu}$, which means that each additional factor of $\gamma_{\mu\nu}$ entering the integral brings an additional power of the coupling constant ℓ_p. This coupling has dimensions of length and thus the theory is non-renormalizable.

Constraints

Another issue to be considered is the diffeomorphism invariance described in Sec. I.B.3. The constraints (I.B.30) it implies in a 3+1 formulation require careful treatment. For example, if one attempts to perform a canonical quantization where the variables $\{h_{ij}\}$ and their conjugate momenta $\{\pi^{ij}\}$ are turned into operators, one must make sure that their commutation relations respect the constraints. This cannot be done in a straightforward manner with the Hamiltonian constraint (I.B.30a) [19].

I.4.3 Quantum cosmology

Related to the issue of applying quantum mechanics to gravity is that of applying it to cosmology. Since this means in principle treating the entire universe as a closed quantum system, the Copenhagen interpretation of quantum mechanics (Sec. I.3.2), with its reliance on an external classical observer performing measurements on the system is inappropriate for quantum cosmology. The early universe presumably contained no sentient observers, and in the vicinity of the Planck time did not even contain any classical subsystems. So a formalism designed for quantum cosmology must be able to describe the quantum mechanics of an entire closed system, and ought to generalize the notion of "measurement" to something defined without reference to an external observer.

I.4.4 Alternate approaches

Before describing the application to quantum gravity of the generalized quantum mechanics formalism, with which this dissertation is concerned, I will briefly mention a few other approaches to quantizing gravity and how they address some of the issues mentioned previously.

String theory

By far the most ambitious of these approaches is string theory [20], which seeks not only to describe gravity quantum mechanically, but to unite it with the other forces of nature as well. Its goal is to describe all particles and forces as a result of the interactions of extended objects. By replacing pointlike particles with one-dimensional strings, string theory introduces a parameter with dimensions of length which provides a natural cutoff scale. This provides a way to overcome the ultraviolet divergence of gravity, and indeed, the theory is perturbatively finite. In addition, since string theory is designed to describe elementary particles as different energy levels in the string spectrum, it has the laudable goal of removing the many arbitrary parameters (masses of quarks and leptons, mixing matrices, *etc.*) found in the standard model of particle physics.

Ashtekar's new variables

This is another prescription whose goal is to avoid the problem of renormalizability, this time by circumventing it altogether and treating gravity non-perturbatively [21]. The idea is to perform a non-perturbative canonical quantization of general relativity. The program avoids some of the constraint problems by working in a new set of variables in which the constraints are more manageable.[9] The spatial metric $\{h_{ij}\}$ and its conjugate momentum $\{\pi^{ij}\}$ are replaced by variables related to the Christoffel symbols (also known as connection coëfficients) $\{\Gamma^{\alpha}_{\beta\gamma}\}$ [*cf.* (I.B.6)], whose conjugate momenta are related to a *tetrad* of four four-vectors $\{E^a_\mu|a=0,1,2,3\}$ which define the spacetime metric via $g_{\mu\nu}=E^a_\mu E^b_\nu\eta_{ab}$ (where $\{\eta_{ab}\}$ is the Minkowski metric $\mathrm{diag}\{-1,1,1,1\}$).

[9] It should be noted that as a fundamentally canonical theory, this approach is susceptible to criticisms related to the problem of time.

Recent formulations [22] have taken advantage of the theory's resemblence to Yang-Mills theory to reformulate it in terms of an analogy to Wilson loops. This "loop representation" makes the fundamental objects of the theory loops and knots, thereby providing a candidate for a more fundamental theory underlying spacetime, just as string theory does.

Euclidean path integrals

This approach [23] applies the path integral quantization methods of field theory to GR, so that the amplitude to make a transition between one metric on a three-surface to another metric on another three-surface is given by a path integral over all possible four-metrics which join the two surfaces. One of the difficulties in performing this path integral is that the integral of $\exp(iS)$ does not converge, but has an oscilatory behavior, and the euclidean path integral approach circumvents this by analytically continuing the metric to one with signature $(+,+,+,+)$ rather than $(-,+,+,+)$.

This is analogous to the sort of continuation performs on a Lorentz metric

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2; \tag{I.4.6}$$

by defining $\tau = it$, the metric takes the form

$$ds^2 = d\tau^2 + dx^2 + dy^2 + dz^2, \tag{I.4.7}$$

which is simply a Euclidean metric. This coördinate transformation is not very useful when t is real, and thus τ pure imaginary, but if one performs calculations for real τ, one can often analytically continue the results back to imaginary τ (and this real t). This is useful for the integral of e^{iS} because of the form the action often takes in the new coördinates. For example, the Klein-Gordon action (I.B.15) on a Minkowski background can be transformed into

$$\begin{aligned} S_{\text{KG}} &= -\frac{1}{2}\int dt\, d^3x[-(\partial_t\varphi)^2 + (\boldsymbol{\nabla}\varphi)^2 + m^2\varphi^2] \\ &= \frac{i}{2}\int d\tau\, d^3x[(\partial_\tau\varphi)^2 + (\boldsymbol{\nabla}\varphi)^2 + m^2\varphi^2] = iS_E \end{aligned} \tag{I.4.8}$$

where the Euclidean action S_E is positive definite for real τ. This means that $e^{iS} = e^{-S_E}$ is exponentially damped rather than oscilliatory as the action moves away from its minimum value. Results calculated in this manner can then be analytically continued back to real t.

The problem with this as a scheme for quantizing gravity is that a general curved metric cannot be analytically continued in this way [14], so its usefulness is limited to situations such as scattering where there is a background spacetime (usually asymptotically flat) on which this phase rotation can be performed.

I.4.5 The decoherence functional for GR

Generalized quantum mechanics is not really a quantization scheme on the same footing as the other approaches just described; it is a way to formulate a quantum-mechanical theory of the universe, whatever the details of that theory might be. As such, it could be used as a framework for one of the approaches discussed in Sec. I.4.4. For example, one might construct the decoherence functional between pairs of histories described by string variables, or use Ashtekar's new variables to produce a decoherence functional rather than a canonical operator description.

As a first application, however, it is worthwhile to write down the decoherence functional for simple general relativity. While we might expect that problems related to renormalization and constraints will still exist, we can see how generalized quantum mechanics addresses the problem of time and the interpretational issues of quantum cosmology without committing to the details of one specific modification of GR.

In addition [24], it seems reasonable that since the fundamental quantum theory of gravity, whatever it is, must reduce to GR in the classical limit, there may be some intermediate regime in which quantum gravitational effects are important, but the differences between the more fundamental theory and GR are not significant. For coarse grainings describing that regime (presumably those which consider averages of fields over regions of Planckian size or larger), we then expect (for suitable initial and final conditions)

$$D_{\text{fundamental}}(\alpha,\alpha') \approx D_{\text{quantum GR}}(\alpha,\alpha'). \tag{I.4.9}$$

Let us then attempt to write down the decoherence functional for general relativity coupled to matter fields, which we write schematically as φ. The sum-over-histories prescription is the most generic, at least in its formal form, so it will be of almost exclusive interest for this dissertation. The decoherence functional is written as

$$D(\alpha,\alpha') = \frac{\sum\limits_{i,j} p_i'' \langle \Phi_i | C_\alpha | \Psi_j \rangle \langle \Phi_i | C_{\alpha'} | \Psi_j \rangle^* p_j'}{\sum\limits_{i,j} p_i'' \left| \langle \Phi_i | C_u | \Psi_j \rangle \right|^2 p_j'}. \tag{I.3.12}$$

The tricky question is how to realize the sum-over-histories prescription (I.3.13) for the matrix elements $\{\langle\Phi|C_\alpha|\Psi\rangle\}$. The straightforward answer would seem to be

$$\langle\Phi|C_\alpha|\Psi\rangle = \Phi^*[\{g''_{\mu\nu}\},\varphi''] \circ \langle\{g''_{\mu\nu}\}\varphi''|C_\alpha|\{g'_{\mu\nu}\}\varphi'\rangle \circ \Psi[\{g'_{\mu\nu}\},\varphi'] \tag{I.4.10a}$$

$$\langle\{g''_{\mu\nu}\}\varphi''|C_\alpha|\{g'_{\mu\nu}\}\varphi'\rangle = \int\limits_{\{g''_{\mu\nu}\}\alpha\{g'_{\mu\nu}\}} \mathcal{D}^6 h \mathcal{D}^6 \pi \mathcal{D}^4 N \mathcal{D}\varphi e^{i(S_{\text{can}}[\{g_{\mu\nu}\}]+S_\varphi[\varphi])}, \tag{I.4.10b}$$

with such niceties as the inner product $\circ$ and the practical realization of the formal sum over histories yet to be specified. However, that is not quite right because of the diffeomorphism invariariance of the theory. First of all, for $\langle\Phi|C_\alpha|\Psi\rangle$ to be diffeomorphism invariant (see Sec. III.6) the wavefunctions must actually depend only upon the spatial metric and not the lapse and shift, so, for example, $\Psi[\{g'_{\mu\nu}\},\varphi']$ is replaced with $\Psi[\{h'_{ij}\},\varphi']$. And secondly, to avoid an infinite factor related to the volume of the gauge group, we must impose a gauge-fixing prescription [25] by including a gauge-fixing delta function $\delta[G]$ and its associated

Fadeev-Popov determinant Δ_G. With these modifications, (I.4.10) becomes

$$\langle\Phi|C_\alpha|\Psi\rangle = \Phi^*[\{h''_{ij}\},\varphi''] \circ \langle\{h''_{ij}\}\varphi''|C_\alpha|\{h'_{ij}\}\varphi'\rangle \circ \Psi[\{h'_{ij}\},\varphi'] \tag{I.4.11a}$$

$$\langle\{h''_{ij}\}\varphi''|C_\alpha|\{h'_{ij}\}\varphi'\rangle = \int_{\{h''_{ij}\}\alpha\{h'_{ij}\}} \mathcal{D}^6 h\mathcal{D}^6\pi\mathcal{D}^4 N\delta[G]\Delta_G[\{g_{\mu\nu}\}]e^{i(S_{\text{can}}[\{g_{\mu\nu}\}]+S_\varphi[\varphi])}. \tag{I.4.11b}$$

There are still many delicacies involved in this path integral, and my approach will be to consider them a few at a time by studying the application of generalized quantum mechanics to simpler theories exhibiting some of the same features.

I.4.6 Toy models

Analogy to the relativistic particle

As discussed in Sec. I.B.3 of Appendix I.B, GR is invariant under diffeomorphisms $x^\mu \to \overline{x}^\mu(\{x^\mu\})$. One subset of the diffeomorphism group is the set of time reparametrizations $(\{x^i\},t) \to (\{x^i\},\overline{t}(t))$. The details of a particular time parametrization are described by the lapse N and shift $\{N^i\}$ from the $3+1$ formulation of GR (see Appendix I.B). The lapse N describes the spacing of the surfaces, while the shift $\{N^i\}$ describes the relation between spatial coördinates $\{x^i\}$ and those on a surface dt into the future.

A much simpler parametrized theory is that of a free relativistic particle in Minkowski space. There the fine-grained histories are paths through spacetime which can be parametrized as functions of a parameter λ, namely $x(\lambda)$. Since a free relativistic particle follows a path of maximal proper time, one expression for its action is

$$S[x] = \int_0^1 d\lambda L - m\int_0^1 d\lambda\frac{d\tau}{d\lambda} = -m\int_0^1 d\lambda\sqrt{-\eta_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}} \tag{I.4.12}$$

This is a difficult action to deal with since it is not quadratic in the velocities $\{\frac{dx^\mu}{d\lambda}\}$. To obtain a quadratic action with the same extrema, one can convert to canonical form. The conjugate momenta are

$$p_\mu = m\eta_{\mu\nu}\frac{dx^\nu/d\lambda}{d\tau/d\lambda}m\frac{dx_\mu/d\lambda}{d\tau/d\lambda}; \tag{I.4.13}$$

the Hamiltonian

$$H = p_\mu\frac{dx^\mu}{d\lambda} - L \tag{I.4.14}$$

vanishes, but there is a constraint

$$\eta^{\mu\nu}p_\mu p^\mu + m^2 = p^2 + m^2 = 0 \tag{I.4.15}$$

obeyed by the momenta, so the canonical action is

$$S_{\text{can}} = \int_0^1 d\lambda\left(p\cdot\frac{dx}{d\lambda} - N\,\frac{p^2+m^2}{2m}\right), \tag{I.4.16}$$

where $N(\lambda)$ is a Lagrange multiplier. The constraint hamiltonian in (I.4.16) has been expressed in a form which makes the action quadratic in the momenta. Varying (I.4.16) with respect to N gives (I.4.15), varying with respect to x gives the equation of motion

$$\frac{dp}{d\lambda} = 0, \tag{I.4.17}$$

and varying with respect to p gives

$$p_\mu = m\frac{dx_\mu/d\lambda}{N}. \tag{I.4.18}$$

Classically, then, the multiplier takes on the value $N = \frac{d\tau}{d\lambda}$.

The analogy between this theory of a relativistic particle and GR is then as follows:

- A reparametrization $\overline{\lambda}(\lambda)$ of the relativistic particle corresponds to a reparametrization $\overline{t}(t)$ of GR.
- The lagrange multiplier N of the relativistic particle theory, which classically obeys $d\tau = Nd\lambda$, corresponds to the lapse function of GR, which obeys, when $\{N^i\}$ and $\{dx^i\}$ both vanish, $d\tau = Ndt$.
- In each theory, the multiplier imposes a quadratic constraint. For the relativistic particle, this is the mass-shell constraint $\frac{1}{2m}(p^2 + m^2) = 0$, while for GR it is the hamiltonian constraint (I.B.30a).

Chapter II is concerned with the generalized quantum mechanics of this theory, particularly with spacetime coarse-grainings of the particle paths.

Analogy to non-Abelian gauge theory

Another subset of diffeomorphism invariance is that in which the time parametrization is unchanged and a spatial diffeomorphism $(\{x^i\}, t) \to \big(\{\overline{x}^j(\{x^i\})\}, t\big)$ is executed. under this spatial diffeomorphism the spatial metric is changed according to

$$\delta h_{ij} = -\,{}^3\nabla_i \delta x_i - {}^3\nabla_j \delta x_j, \tag{I.4.19}$$

along with changes in the momenta, lapse and shift. These spatial diffeomorphisms are analogous to the gauge transformations in a non-Abelian gauge theory with canonical action

$$S_{\text{can}} = \int d^4x(\boldsymbol{\pi}_a \cdot \dot{\mathbf{A}}_a - \mathcal{H}[\mathbf{A}, \boldsymbol{\pi}] - \varphi_a K_a). \tag{I.4.20}$$

The infinitesimal gauge transformations under which this theory is invariant change the vector potential acording to

$$\delta \mathbf{A}_a = -\boldsymbol{\nabla}\delta\Lambda_a - g f_{ab}^c \mathbf{A}_c \delta\Lambda_b = (-\mathbf{D}\delta\Lambda)_a\,, \tag{I.4.21}$$

along with changes in the momenta and the scalar potential φ.

The analogy between GR and a non-Abelian gauge theory (NAGT) is thus

- The gauge transformations (I.4.21) in a NAGT correspond to spatial diffeomorphisms (I.4.19) in GR.
- The scalar potential φ in a NAGT is a lagrange multiplier, corresponding to the shift vector $\{N^i\}$ in 3+1 GR.
- The scalar potential enforces the Gauss's law constraint $K = \mathbf{D} \cdot \boldsymbol{\pi}$ in a NAGT, which corresponds to the momentum constraint (I.B.30b) of GR.

I.5 Outline

The plan for the rest of this dissertation is as follows:

I.5.1 Chapter II: Spacetime coarse grainings in relativistic particle QM

In Chapter II, I use generalized quantum mechanics to predict probabilities for sets of spacetime alternatives (see Sec. I.2) in the quantum mechanics of a single relativistic particle (see Sec. I.4.6). In earlier work, Yamada and Takagi [2] considered certain spacetime coarse grainings in nonrelativistic quatum mechanics, and exhibited such coarse grainings which decohered for initial conditions with particular symmetry properties. In Chapter II, I examine similar coarse grainings in the case of the relativistic particle, and find that some, but not all, of them decohere and allow the consistent assignment of probabilities.

While these alternatives are extremely simple and idealized, they do provide an example of how generalized quantum mechanics can make predict probabilities for non-trivial sets of spacetime alternatives in this simple relativistic theory.

I.5.2 Chapter III: Non-Abelian gauge theories

Gauge invariance is the primary focus of Chapter III, which describes a sum-over-histories generalized quantum mechanics of a non-Abelian gauge theory (NAGT) with an arbitrary gauge group and no matter. [The motivation is not to apply this generalized quantum mechanics directly to the strong or weak interaction, but to learn more about the consequences of the gauge symmetry with an eye towards applying these lessons to the gauge (diffeomorphism) symmetry of general relativity, according to the analogy described in Sec. I.4.6.] In a NAGT, because of the gauge symmetry, there are a number of choices to be made in the formulation of the quantum mechanics itself. A frequently-used tactic in quantizing gauge theories is to impose *a priori* the constraints corresponding to the gauge symmetries, and quantize only the variables in the so-called physical subspace. Since it generally difficult to isolate the physical subspace, and because giving the constraints special status breaks the manifest Lorentz invariance of the theory, I instead take the more general approach of performing a sum-over-histories quantization of the entire theory and observing in which ways the constraints manifest themselves. This is done with coarse grainings which ask physical questions corresponding to the constraints. For instance, one can coarse grain by the value of the covariant divergence of the non-Abelian electric field. I use a sum-over-histories generalized quantum mechanics to make such predictions, and compare the results to those of quantum-mechanical descriptions of gauge theories in which the constraints are imposed by hand. When the constraint is expressed in the natural way in terms of phase space variables, it is seen to hold, both through a simple formal argument and via a more careful description of the Lorentzian path integral as defined on a spacetime lattice. However, if the constraint is defined using only configuration space variables (the scalar and vector potential, as opposed to the momentum conjugate to the latter), the situation is more complicated. For some coarse grainings, I recover the known result from an abelian theory, that coarse grainings by values of the constraint either predict its vanishing or fail to decohere. However, sets of alternatives defined in terms of a more complicated quantity in the abelian case are exhibited where definite predictions can be made which disagree with the assumption that the constraints vanish. Despite this seeming failure, the configuration space approach is appealing because, as is shown in Sec. III.6, it can be formulated in a manifestly Lorentz-invariant way. In this formulation, the constraints obeyed by the wave functionals are directly related to the surfaces in spacetime on which those wave functionals are defined.

The methods for inducing decoherence in Chapters II and III are somewhat artificial. In Chapter II, some sets of alternatives decohere because symmetries of the initial states cause the interference terms to cancel. In Chapter III, with the exception of Sec. III.5.4, all the decohering sets of alternatives have the property that the probability is one for one alternative and zero for the others. In this sense they are more identities than quantum mechanical predictions. The process which is believed to result in most real-world decoherence [12] involves a division into a "system" of physical interest coupled to an "environment" which carries away phase information, causing coarse grainings describing only "system" variables to decohere.

I.5.3 Chapter IV: Modelling the decoherence of spacetime

Chapter IV poses the question of when such physical decoherence occurs in the gravitational field, a phenomenon known as the decoherence of spacetime. Previous work [26, 27] has used an additional field to obtain decoherence of the gravitational field in cosmological models. In Chapter IV I address the question of whether decoherence can be induced in the gravitational field itself by separating the field into short- and long-wavelength modes and allowing coarse grainings to refer only to the long-wavelength modes. In this way one may model the intuitively reasonable effect that if we examine the gravitational field on scales long compared to the Planck length, quantum gravitational interference will become unimportant, and one will be able to make predictions for alternatives defined on those scales. (And presumably, at large enough scales, those predictions will corrsepond to those of classical GR.) For a toy model in which the perturbative action for gravity is mimicked by a scalar field, I demonstrate such decoherence for coarse grainings which have certain properties. Namely, the short-wavelength modes are taken to be in a thermal state, whose temperature is higher than that corresponding to the length scale dividing "short" and "long" wavelengths, and the modes defining the coarse graining have temporal frequencies lower then their spatial frequencies. Under these circumstances, decoherence will occur when the alternatives are sufficiently coarse-grained. The question of decoherence in other regimes cannot be addressed by the perturbative analysis in Chapter IV.

I.A Notation and conventions

I.A.1 Units

Except when it is convenient to show factors of c or $\hbar$ explicitly, I will be using the system of units in which $c = 1 = \hbar$, so for instance $1\,\mathrm{cm} = 1\,\mathrm{cm} = 3.3 \times 10^{-11}\,\mathrm{s} = 5.1 \times 10^{10}\,\mathrm{MeV}^{-1}$. Factors of Newton's constant, will, however, always be explicitly included, and its value is $G = m_p^{-2} = \ell_p^2$.

I.A.2 Sign conventions

In general, I will use the sign conventions of [14] for relativity. In particular, the spacetime metric has one negative eigenvalue and three positive ones, so the Minkowski metric is

$$\{\eta_{\mu\nu}\} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{I.A.1}$$

I.A.3 Vectors and indices

I.A.4 Indices

I will represent spacetime indices with greek letters $(\mu, \nu, \ldots)$, and such indices take on values $0, \ldots, D$, where D is the number of spatial dimensions, usually 3. Spatial indicies are represented by latin letters from the middle of the alphabet $(i, j, \ldots)$ and run from 1 to D. Except in Chapter IV, latin indices from the beginning of the alphabet $(a, b, \ldots)$ are "other" indices, either labelling the vectors in a tetrad or generator indices in a gauge theory. Their position as superscripts or subscripts is arbitrary, as opposed to spacetime indices which are raised and lowered with the metric $\{g_{\mu\nu}\}$ and spatial indices which are raised and lowered with the metric $\{h_{ij}\}$. In Chapter IV, latin indices $a, b, \ldots$ are part of the abstract index notation described in Sec. IV.3.1. All of these indices obey the Einstein summation convention in which repeated indices are summed over their entire range. For any other indices (such as **M** and **N** from Chapter IV), summation occurs only where explicitly stated.

Vectors and tensors

General D or $D+1$ vectors on curved space(time) are written as sets of components $\{v^i\}$ or $\{v^\mu\}$. Spatial vectors on flat Euclidean space are also written in boldface, as in $\mathbf{E}$. Assorted vectors and matrices, the interpretation of which should be apparent from context, are written as unadorned letters v. This includes vectors in flat $D + 1$ Minkowski space.

When an object such as k or ∇ is to be interpreted as a scalar rather than a four-vector, I will endeavor to point this out.

I.A.5 Functions and functionals

The argument of a function will generally be written in parentheses, $\mathbf{A}'(\mathbf{x})$; the argument of a functional will be written in square brackets, $\Psi[\mathbf{A}']$. When an object is a function of some variables and a functional of others, I will use a mixed-parentheses notation. For example, $\Psi[\mathbf{A}'; t')$ is a functional of $\mathbf{A}'(\mathbf{x})$ and a function of t'. In the case of propagators, I will divide the arguments with a semicolon on each side of the vertical bar, as in $\mathcal{G}\,[\mathbf{A}''; t''|\, \mathbf{A}'; t)$. Finally, wherever possible, I will try to distinguish between functions over spacetime x and functions over space $\mathbf{x}$ by writing the latter with a prime, an index, or the like. So $S[\varphi]$ is a functional of spacetime field configurations $\varphi(x)$ while $\Phi[\varphi']$ is a functional of spatial field configurations $\varphi'(\mathbf{x})$.

Miscellaneous

The Heavyside step function is given by

$$\Theta(x) = \begin{cases} 1, & x > 0 \\ 0, & x < 0. \end{cases} \tag{I.A.2}$$

I.B General Relativity

Gravitational phenomena at scales ranging from small macroscopic bodies falling on the Earth to the motion of clusters of galaxies are accurately described by Einstein's general theory of relativity (GR). This replaced the Newtonian theory of gravity, in which the gravitational potential φ is an external field which acts upon objects according to

$$\mathbf{F} \equiv m\frac{d^2\mathbf{x}}{dt^2} = -m\boldsymbol{\nabla}\varphi \tag{I.B.1a}$$

and is generated by the mass density ρ of matter:

$$\boldsymbol{\nabla}^2\varphi = 4\pi G\rho, \tag{I.B.1b}$$

In contrast, the action of gravity in GR is not some external force field, but an alteration in the geometry of spacetime itself. This formulation has built into it what is a seeming coïncidence in the Newtonian theory, namely that the inertial mass m appearing on the left-hand side of (I.B.1a) is the same as the parameter m on the right-hand side which describes how an object couples to the gravitational field, and thus the parameters cancel out. This is known as the *equivalence principle*, and means that any object, regardless of its mass, will follow the same trajectory in a gravitational field. (Contrast this with the case of electromagnetic fields, where the charge-to-mass ration of the object comes into play.) This allows the effect of gravity to be described as a property of the spacetime in which all objects move.

Among other things, the equivalence principle also means that if an entire system is falling in a gravitational field, it is impossible, by conducting experiments within that system, to detect the gravitational field. It is only possible to detect inhomogeneities in the field, if the system is large enough that the difference between the fields at both extremes becomes detectable. Then there will be a relative acceleration between different elements of the system. This also has a mathematical expression in general relativity.

I.B.1 The geometry of spacetime

The spacetime metric

The geometry of spacetime is described by the metric $g_{\mu\nu}$, which is used to define the spacetime interval ds^2 between two points by

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu. \tag{I.B.2}$$

In flat (Minkowski) spacetime, it is possible to choose a global set of coördinates $\{x^0, x^1, x^2, x^3\}$ such that

$$ds^2 = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2, \tag{I.B.3}$$

but this is only true locally in a general curved spacetime. To see this, consider a general metric $\{g_{\overline{\mu\nu}}(x)\}$, which is a four-by-four symmetric matrix, and thus has ten independent components which are functions of the position x in spacetime. Under a change of coördinates from $\{x^\mu\}$ to $\{x^{\overline{\mu}}\}$, application of the chain rule of multivariable calculus to (I.B.2) and $ds^2 = g_{\overline{\mu\nu}} dx^{\overline{\mu}} dx^{\overline{\nu}}$ gives the transformation law

$$g_{\mu\nu} = g_{\overline{\mu\nu}} \frac{\partial x^{\overline{\mu}}}{\partial x^\mu} \frac{\partial x^{\overline{\nu}}}{\partial x^\nu} \tag{I.B.4}$$

Thus we see that we have sixteen first derivatives $\{\partial x^{\overline{\mu}}/\partial x^\mu\}$ which can be chosen to set the ten[10] independent components $\{g_{\mu\nu}\}$ to their flat-space values $g_{\mu\nu} = \eta_{\mu\nu}$ (where $\eta_{00} = -1$, $\eta_{01} = 0$, $\eta_{02} = 0$, $\eta_{03} = 0$, $\eta_{11} = 1$, *etc.*) at a single point, with six degrees of freedom to spare. If we move on and try to set the first derivatives $\left\{\frac{\partial g_{\mu\nu}}{\partial x^\lambda}\right\}$ to their flat-space values, namely zero, this gives us $10 \times 4 = 40$ additional values to fix, and we have at our disposal the $4 \times 10 = 40$ second derivatives $\left\{\frac{\partial^2 x^{\overline{\mu}}}{\partial x^\mu \partial x^\nu}\right\}$ describing the coördinate transformation. Finally we lose the ability to fix things completely when we consider the second derivatives $\left\{\frac{\partial^2 g_{\mu\nu}}{\partial x^\lambda \partial x^\sigma}\right\}$ of the metric. There are $10 \times 10 = 100$ of these, and only $4 \times 20 = 80$ third derivatives $\left\{\frac{\partial^3 x^{\overline{\mu}}}{\partial x^\mu \partial x^\nu \partial x^\lambda}\right\}$. This means that, at least by this simple counting argument, there should be $100 - 80 = 20$ independent components $\left\{\frac{\partial^2 g_{\mu\nu}}{\partial x^\lambda \partial x^\sigma}\right\}$ which cannot be set to zero by a coördinate transformation.

[10] A four-by-four symmetric matrix has $\binom{4}{1} = 4$ diagonal components plus $\binom{4}{2} = 6$ off-diagonal components. A fully-symmetric three-index object like $dx^\mu dx^\nu dx^\lambda$ has 4 "diagonal" components, $2! \times \binom{4}{2} = 12$ components where two indices are the same and the third is different, and $\binom{4}{3} = 4$ components where all four indices are different, for a total of 20 independent components.

Transformations, tensors and the invariant description of curvature

To see what these twenty combinations of $\left\{\frac{\partial^2 g_{\mu\nu}}{\partial x^\lambda \partial x^\sigma}\right\}$ are requires just a bit of geometry. Equation (I.B.4) tells us that the metric behaves as a second rank covariant[11] tensor under changes of coördinates; likewise, a spacetime displacement dx is a contravariant four-vector (first rank tensor) obeying transformation the law $dx^\mu = \frac{\partial x^\mu}{\partial x^{\overline{\mu}}} dx^{\overline{\mu}}$. These are all examples of the general tensor transformation law

$$T^{\nu_1\cdots\nu_n}_{\mu_1\cdots\mu_m} = T^{\overline{\nu}_1\cdots\overline{\nu}_n}_{\overline{\mu}_1\cdots\overline{\mu}_m} \frac{\partial x^{\overline{\mu}_1}}{\partial x^{\mu_1}} \cdots \frac{\partial x^{\overline{\mu}_m}}{\partial x^{\mu_m}} \frac{\partial x^{\nu_1}}{\partial x^{\overline{\nu}_1}} \cdots \frac{\partial x^{\nu_n}}{\partial x^{\overline{\nu}_n}} \tag{I.B.5}$$

which is the generalization of the familiar Lorentz transformation law, with the Lorentz transformation matrix $\{\Lambda^{\overline{\mu}}_{\mu}\}$ replaced by $\left\{\frac{\partial x^{\overline{\mu}}}{\partial x^\mu}\right\}$. The ordinary gradient $\partial_\lambda T^{\nu_1\cdots\nu_n}_{\mu_1\cdots\mu_m} = \frac{\partial T^{\nu_1\cdots\nu_n}_{\mu_1\cdots\mu_m}}{\partial x^\lambda}$ of a tensor does not have these transformation properties, as must be the case in light of our demonstration that $\partial_\lambda g_{\mu\nu}$ can always be made to vanish at a point in some coördinate system, while this is clearly not true for an arbitrary metric in an arbitrary coördinate system. (Since the transformation law (I.B.5) is homogeneous and linear, a tensor which vanishes in one coördinate system must vanish in every one.) Instead, the operation must be converted to a covariant derivative ∇_λ which is corrected from ∂_λ as follows:

$$\nabla_\lambda T^{\nu_1\cdots\nu_n}_{\mu_1\cdots\mu_m} = \partial_\lambda T^{\nu_1\cdots\nu_n}_{\mu_1\cdots\mu_m} + \sum_{j=1}^{n} \Gamma^{\nu_j}_{\lambda\sigma} T^{\nu_1\cdots\sigma\cdots\nu_n}_{\mu_1\cdots\mu_m} - \sum_{i=1}^{n} \Gamma^{\sigma}_{\lambda\mu_i} T^{\nu_1\cdots\nu_n}_{\mu_1\cdots\sigma\cdots\mu_m} \tag{I.B.6}$$

The quantity $\nabla_\lambda T^{\nu_1\cdots\nu_n}_{\mu_1\cdots\mu_m}$ *will* transform as a tensor under coördinate changes, and can thus be used to contruct coördinate-invariant physical laws. The *Christoffel symbols* $\{\Gamma^\lambda_{\mu\nu}\}$ which accomplish this are given by

$$\Gamma^\lambda_{\mu\nu} = \frac{g^{\lambda\sigma}}{2} \left(\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}\right), \tag{I.B.6a}$$

where $g^{\mu\nu}$ are the components of the inverse metric defined by $g^{\mu\lambda} g_{\lambda\nu} = \delta^\mu_\nu$. The Christoffel symbols $\{\Gamma^\lambda_{\mu\nu}\}$ are not the components of a tensor (since they can always be made to vanish at a point by a suitable coördinate transformation), and are thus a property of the coördinate system and not of the spacetime geometry alone. In fact, the statement that there is always a coördinate system for which, at a particular point, $g_{\mu\nu} = \eta_{\mu\nu}$ and $\Gamma^\mu_{\nu\lambda} = 0$ is another statement of the equivalence principle.

The tensorial object describing the curvature of the geometry is given by considering the fact that covariant derivatives, unlike ordinary ones, do not commute, and it turns out [14] that given any tensor $\{T^{\nu_1\cdots\nu_n}_{\mu_1\cdots\mu_m}\}$

$$\nabla_\lambda \nabla_\sigma T^{\nu_1\cdots\nu_n}_{\mu_1\cdots\mu_m} - \nabla_\sigma \nabla_\lambda T^{\nu_1\cdots\nu_n}_{\mu_1\cdots\mu_m} = -\sum_{j=1}^{n} R_{\lambda\sigma\rho}{}^{\nu_j} T^{\nu_1\cdots\rho\cdots\nu_n}_{\mu_1\cdots\mu_m} + \sum_{i=1}^{n} R_{\lambda\sigma\mu_i}{}^{\rho} T^{\nu_1\cdots\nu_n}_{\mu_1\cdots\rho\cdots\mu_m} \tag{I.B.7}$$

[11] In the language of differential geometry, a subscripted index is known as a covariant index, while a superscripted index is called contravariant. The tensor transformation laws are designed so that if all of the contravariant and covariant indices are paired off and summed over (as in, for example, $g_{\mu\nu} u^\mu u^\nu$), the resulting object is a scalar, *i.e.*, is unchanged by a coördinate transformation.

$\{R_{\mu\nu\lambda}{}^{\sigma}\}$ are the components of a tensor known as the Riemann curvature tensor, and are given in a particular coördinate system by

$$R_{\mu\nu\lambda}{}^{\sigma} = \partial_\nu \Gamma^{\sigma}_{\mu\lambda} - \partial_\mu \Gamma^{\sigma}_{\nu\lambda} + \Gamma^{\rho}_{\mu\lambda}\Gamma^{\sigma}_{\rho\nu} - \Gamma^{\rho}_{\nu\lambda}\Gamma^{\sigma}_{\rho\mu} \tag{I.B.8}$$

These combinations of second derivatives of the metric (and there turn out to be twenty independent quantities among them) are what describe the curvature of the spacetime. They have, among others, the following symmetries:

$$R_{\mu\nu\lambda}{}^{\sigma} = -R_{\nu\mu\lambda}{}^{\sigma} = R_{\lambda}{}^{\sigma}{}_{\mu\nu}, \tag{I.B.9}$$

where we have used the metric and its inverse to "raise" and "lower" indices ($R_{\lambda}{}^{\sigma}{}_{\mu\nu} = g^{\sigma\rho} g_{\nu\xi} R_{\lambda\rho\mu}{}^{\xi}$). They are clearly zero for flat spacetime described by the metric (I.B.3) (which has vanishing Christoffel symbols), but because of the tensorial nature of the Riemann tensor, they also vanish for any description of the same geometry in another coördinate system, for example spherical polar coördinates, in which

$$ds^2 = -dt^2 + dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2. \tag{I.B.10}$$

In this coördinate system, some of the first derivatives of the metric coëfficients are non-zero (for example, $\partial_r g_{\theta\theta} = 2r$), but all of the components of $R_{\mu\nu\lambda}{}^{\sigma}$ still vanish, showing that the spacetime is still flat.

The motion of particles in curved spacetime

This, then, is how one describes the geometry of spacetime mathematically. Its effects on the motion of objects are as follows: A particle's trajectory is given by some path $x(\lambda)$ through spacetime, where λ parametrizes the path. A convenient choice of parameter is the proper time τ, given by $d\tau^2 = -ds^2$ along the path. Given that τ is an invariant (scalar) quantity, the four-velocity with components $u^\mu = \frac{dx^\mu}{d\tau}$ is a contravariant four-vector, and by the chain rule $\frac{d}{d\tau} = u^\mu \partial_\mu$. Now, a particle moving under the influence of no external forces in non-relativistic mechanics obeys

$$\frac{d^2\mathbf{x}}{dt^2} = 0. \tag{I.B.11}$$

Keeping in mind that general relativity is constructed so that the action of gravity is not an explicit force, but a consequence of geometry, the motion of an object moving under the influence only of gravity must obey a similar, relativistic equation, as dictated by the equivalence principle. Replacing the time with proper time and the three-vector $\mathbf{dx}$ with the four-vector dx, the left hand side of (I.B.11) becomes $\frac{d^2x^\mu}{d\tau^2}$. This is not quite a contravariant vector, since it is equal to $\frac{du^\mu}{d\tau} = u^\nu \partial_\nu u^\mu$. The obvious modification is to replace the ordinary gradient with a covariant one to give

$$u^\nu \nabla_\nu u^\mu = \frac{d^2x^\mu}{d\tau^2} + \Gamma^{\mu}_{\nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} = 0. \tag{I.B.12}$$

This is called the geodesic equation, and it describes the motion of an object effected only by the spacetime geometry, just as (I.B.1a) describes the motion of an object effected only by the gravitational field. [The effect of external forces is be described by adding terms to the right hand side of (I.B.12), just as the effect of forces other than gravity would be described by adding terms to the right hand side of (I.B.1a).]

The geodesic equation also allows one to calculate the relative acceleration described in the preamble to this section. If two particles are initially moving along parallel trajectories with a four-velocity $\{u^\mu\}$ and separated by a small displacement $\{n^\mu\}$, their separation will change with time according to the *equation of geodesic deviation* [17]

$$u^n u \nabla_\nu u^\lambda \nabla_\lambda n^\mu = -R_{\nu\lambda\sigma}{}^\mu u^\nu n^\lambda u^\sigma \tag{I.B.13}$$

I.B.2 How matter generates curvature

The stress-energy tensor

With (I.B.12) describing the effect of the spacetime geometry on the motion of particles, taking the place of the non-relativistic, Newtonian formula (I.B.1a), the other half of general relativity describes the way in which matter generates spacetime curvature, which takes the place of (I.B.1b). The relativistic quantity corresponding to the mass density ρ appearing in (I.B.1b) is the stress-energy tensor $\{T_{\mu\nu}\}$. In Minkowski space, the component T_{00} is equal to the energy density ρ, the mixed components $\{T_{i0} = T_{0i}\}$ are equal to the momentum density in the ith spatial direction, or the flux of energy across a surface oriented in that direction, and the spatial components $\{T_{ij}\}$ are the components of stresses. Given a matter field with action S_M, the stress-energy is defined[12] as the functional derivative of S_M with respect to the metric:

$$T_{\mu\nu} = -\frac{2}{\sqrt{|g|}}\frac{\mathcal{D}S_M}{\mathcal{D}g^{\mu\nu}}. \tag{I.B.14}$$

For example, a Klein-Gordon scalar field φ has an action[13]

$$S_{\text{KG}} = -\frac{1}{2}\int \sqrt{|g|}\, d^4x [g^{\mu\nu}(\nabla_\mu\varphi)(\nabla_\nu\varphi) + m^2\varphi^2]; \tag{I.B.15}$$

Using the identity

$$\frac{\mathcal{D}}{\mathcal{D}g^{\mu\nu}}\int \sqrt{|g|}\, d^4x \mathcal{L}(x) = \int \sqrt{|g|}\, d^4x \frac{\mathcal{D}\mathcal{L}(x)}{\mathcal{D}g^{\mu\nu}} - \frac{\sqrt{|g|}}{2}\mathcal{L}g_{\mu\nu} \tag{I.B.16}$$

one can calculate

$$T^{\text{KG}}_{\mu\nu} = (\nabla_\mu\varphi)(\nabla_\nu\varphi) - \frac{g_{\mu\nu}}{2}\left[g^{\lambda\sigma}(\nabla_\lambda\varphi)(\nabla_\sigma\varphi) + m^2\varphi^2\right] \tag{I.B.17}$$

for the stress-energy tensor.

[12] There are actually some subtleties in the relationship between this definition of the stress-energy tensor and the perhaps more familiar definition as a Noether current. See [28] for details.

[13] Since $\sqrt{|g|}\, d^4x = \sqrt{|\det\{g_{\mu\nu}\}|}\, dx^0 dx^1 dx^2 dx^3$ is the invariant volume element, the action (I.B.15) is invariant.

Likewise, the action for an electromagnetic field (not including the source terms) is

$$S_{\text{EM}} = -\frac{1}{4}\int \sqrt{|g|}\, d^4x F_{\mu\nu}F^{\mu\nu} = -\frac{1}{4}\int \sqrt{|g|}\, d^4x\, g^{\mu\nu}g^{\lambda\sigma}F_{\mu\lambda}F_{\nu\sigma} \tag{I.B.18}$$

which leads to the stress-energy tensor

$$T^{\text{EM}}_{\mu\nu} = g^{\lambda\sigma}F_{\mu\lambda}F_{\nu\sigma} - \frac{g_{\mu\nu}}{4}F_{\lambda\sigma}F^{\lambda\sigma}. \tag{I.B.19}$$

The gravitational action

The spacetime geometry is coupled to the stress-energy tensor via an action principle. Requiring the total action $S = S_G + S_M$ to be stationary under changes of the metric gives

$$\frac{2}{\sqrt{|g|}}\frac{\mathcal{D}S_G}{\mathcal{D}g^{\mu\nu}} = T_{\mu\nu}, \tag{I.B.20}$$

which is the GR counterpart of (I.B.1b). If we wish to construct the gravitational action S_G out of the spacetime geometry alone, the only tensors at our disposal are $\{g_{\mu\nu}\}$ and $\{R_{\mu\nu\lambda}{}^{d}\}$. The simplest scalar which can be created from these two is just a constant Λ, but this does not lead to an interesting theory by itself. To create a scalar which is linear in the curvature tensor, we "contract" it with the metric to get first the Ricci tensor $R_{\mu\nu} = R_{\mu\lambda\nu}{}^{\lambda}$ and finally the curvature scalar $R = g^{\mu\nu}R_{\mu\nu}$. The simplest non-trivial action for the gravitational field is thus

$$S_G = \frac{1}{16\pi G}\int \sqrt{|g|}\, d^4x(R - 2\Lambda) \tag{I.B.21}$$

which leads (ignoring surface terms in the variation) to the Einstein equation

$$R_{\mu\nu} - \frac{g_{\mu\nu}}{2}R + \Lambda g_{\mu\nu} = 8\pi G\, T_{\mu\nu}. \tag{I.B.22}$$

This replaces (I.B.1b), and the effects of the theory reduce to those of Newtonian gravity for small curvature [17].

The constant Λ is known as the cosmological constant, and has had a long and checkered history [29], but as current experimental limits [30] set

$$|\Lambda| \lesssim 5 \times 10^{-56}\,\text{cm}^{-2}, \tag{I.B.23}$$

it will suffice to take $\Lambda = 0$ for the purposes of this dissertation.

I.B.3 Features of the gravitational action

Diffeomorphism invariance

Having briefly described GR, I will now discuss one of its salient features, namely its invariance under changes in coördinate system.[14] This would seem to be little more than

[14] I will describe this transformation as a change of coördinates, *i.e.*, a passive transformation in which a given point in spacetime is relabelled from $\{x^\mu\}$ to $\{x^\mu + \delta x^\mu\}$. It could, of course, also be described in terms of an active diffeomorphism which moves the point originally labelled by $\{x^\mu\}$ to a new location, which is the point labelled in the original coördinates by $\{x^\mu - \delta x^\mu\}$.

a restatement of the fact that since the entire theory is expressed in covariant language, the predictions of the theory are unchanged when the metric is transformed according to (I.B.4). However, there is a subtle, if familiar, point to be taken into consideration, namely that since the left- and right-hand sides of (I.B.4) are supposed to be evaluated at the same point in spacetime, the values of the coördinates in the arguments of $g_{\mu\nu}(\{x^\lambda\})$ and $g_{\overline{\mu\nu}}(\{x^{\overline{\lambda}}\})$ will not be the same. Thus the effect of an infinitesimal coördinate change[15] $x^\mu \to \overline{x}^\mu = x^\mu + \delta x^\mu$ will be, to first order in δx^μ

$$\begin{aligned}\delta g_{\mu\nu} &= \overline{g}_{\mu\nu}(\{x^\lambda\}) - g_{\mu\nu}(\{x^\lambda\}) \\ &= \overline{g}_{\mu\nu}(\{\overline{x}^\lambda\}) + \delta x^\lambda \partial_\lambda g_{\mu\nu} - \overline{g}_{\sigma\rho}(\{\overline{x}^\lambda\})(\delta^\sigma_\mu + \partial_\mu \delta x^\sigma)(\delta^\rho_\nu + \partial_\nu \delta x^\rho) \\ &= \delta x^\lambda \partial_\lambda g_{\mu\nu} - g_{\lambda\nu}\partial_\mu \delta x^\lambda - g_{\mu\lambda}\partial_\nu \delta x^\lambda = -\nabla_\mu \delta x_\nu - \nabla_\nu \delta x_\mu;\end{aligned} \tag{I.B.24}$$

it is under these infinitesimal changes of the function $g_{\mu\nu}(\{x^\lambda\})$ that the action (I.B.21) is invariant.

3+1 and Hamiltonian formulation

The nature of diffeomorphism invariance, and of the conservation laws to which it leads, is further elucidated by dividing the four spacetime coördinates $\{x^\mu\}$ into three "space" coördinates $\{x^i\}$ and one "time" coördinate t. The ten metric functions $\{g_{\mu\nu}(\{x^\mu\})\}$ are replaced by the lapse function $N(t, \{x^i\})$, the three components $\{N^i(t, \{x^i\})\}$ of the shift vector and the six independent components $\{h_{ij}(t, \{x^i\})\}$ of the spatial metric on a constant-t surface. The metric written in terms of these variables is (see Fig. I.2 for a physical interpretation)

$$ds^2 = -N^2 dt^2 + h_{ij}(N^i dt + dx^i)(N^j dt + dx^j). \tag{I.B.25}$$

In terms of these variables, the gravitational action becomes [14], up to a surface term,

$$S_G = \frac{1}{16\pi G}\int dt d^3x\, N\sqrt{h}({}^3R + K^{ij}K_{ij} - K^2) = \int dt L_G, \tag{I.B.26}$$

where ${}^3R$ is the scalar curvature of the surface of constant time, calculated from the three-metric $\{h_{ij}\}$ just as the scalar curvature R is calculated from the four-metric $g_{\mu\nu}$ using (I.B.8), K_{ij} is the extrinsic curvature (defining $\dot{h}_{ij} = \partial_t h_{ij}$)

$$K_{ij} = \frac{1}{2N}\left(\dot{h}_{ij} - {}^3\nabla_i N_j - {}^3\nabla_j N_i\right) \tag{I.B.27}$$

of a constant-time surface, and ${}^3\nabla_i$ is the covariant derivative on the three-surface constructed from the metric $\{h_{ij}\}$. Spatial indices are "raised" using the inverse three-metric

[15] It is now useful to change our notation for the new coördinates from $x^{\overline{\mu}}$ to $\overline{x}^\mu$. This is because in an infinitesimal transformation it is useful to compare $\overline{x}^\mu$ to x^μ. For general transformations, there is no correspondence between components $x^{\overline{\mu}}$ and x^μ, so it is more important to stress that, for example $g_{\overline{\mu\nu}}$ and $g_{\mu\nu}$ are different component realizations of the same physical object.

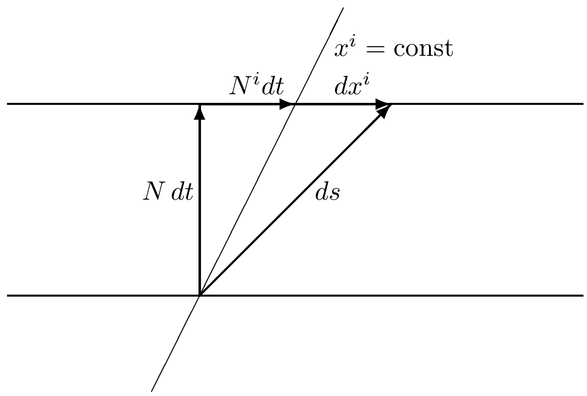


Figure I.2: Dividing a spacetime displacement $\{dx^\mu\}$ into a spatial displacement $\{dx^i\}$ and a time displacement dt. The spacetime displacement ds consists of a timelike displacement $N dt$ where the lapse function N quantifies the separation of two “consecutive” time slices t and $t + dt$ and a spacelike displacement $N^i dt + dx^i$ where the shift vector N^i describes how far the spatial coördinates have “moved” between the two time slices.

$\{h^{ij}\}$ (satisfying $h^{ik}h_{kj} = \delta^k_j$) to give $K^{ij} = h^{ik}h^{j\ell}K_{k\ell}$, and $K = h^{ij}K_{ij}$ is the trace of the extrinsic curvature.

To convert this Lagrangian formulation to a Hamiltonian one, we determine the canonical momenta conjugate to $\{h_{ij}\}$ by[16]

$$\pi^{ij} = 16\pi G \frac{\mathcal{D}L}{\mathcal{D}\dot{h}_{ij}} = \sqrt{h}(K^{ij} - Kh^{ij}); \tag{I.B.28}$$

since the time derivatives $\dot{N}$ and $\{\dot{N}^i\}$ do not appear in the Lagrangian, N and $\{N^i\}$ have no conjugate momenta (they are Lagrange multipliers), and the gravitational action can be rewritten in the canonical form

$$S_{\text{can}} = \frac{1}{16\pi G}\int dt d^3x \left(\pi^{ij}\dot{h}_{ij} - NH - N^i H_i\right) \tag{I.B.29}$$

where

$$H = \frac{1}{2\sqrt{h}}(h_{ik}h_{j\ell} + h_{i\ell}h_{jk} - h_{ij}h_{k\ell})\pi^{ij}\pi^{k\ell} - {}^3R\sqrt{h} \tag{I.B.30a}$$

and

$$H_i = 2\,{}^3\nabla_j \pi^j_i. \tag{I.B.30b}$$

(${}^3\nabla_i$ is the covariant derivative constructed from the three-metric $\{h_{ij}\}$.) Variation of the canonical action with respect to N and $\{N^i\}$ produces the constraints $H = 0$ and $H_i = 0$.

[16]Note that the object with components $\{h_{ij}/\sqrt{h}\}$ transforms as a tensor under changes of the spatial coördinates $\{x^i\}$.

These constraints are used as inspiration for the toy models in the next two chapters: the mass-shell constraint of the relativistic particle action considered in Chapter II is analogous to the Hamiltonian constraint (I.B.30a) while the Gauss's law constraint of the non-Abelian gauge theory action in Chapter III is similar to the momentum constraint (I.B.30b).

Chapter II

Spacetime Alternatives for the Relativistic Particle

II.1 Introduction

One of the reasons why one expects a standard quantum mechanics, described by states on a spacelike surface, to be inadequate to describe quantum gravity is that the notion of "spacelike" should be ill-defined in a theory where the metric itself is behaving quantum mechanically. Standard quantum mechanics makes reference to spacelike surfaces not only in its description of the state of the system "at a moment of time", but also in the very alternatives for which it makes predictions. A theory which predicts spacetime probabilities, such as the probability that a particle passes through an extended region of spacetime during its trajectory, can thus be thought of as one step on the road towards a quantum theory of gravity. Spacetime alternatives in nonrelativistic quantum mechanics have been considered in the past by Feynman [4], Yamada and Takagi [2], and Hartle [1].

This chapter considers spacetime alternatives for the quantum mechanics of a free relativistic particle. This is not meant as a quantum theory of actual relativistic particles (which are described by quantum field theory) but rather as a toy model for quantum cosmology. As described in Sec. I.4.6, the reparametrization invariance of the relativistic particle mimics part of the diffeomorphism invariance of GR and makes this a good toy theory.

This chapter calculates the decoherence functionals for some simple coarse grainings according to the following steps: Section II.2 describes the construction of the decoherence functional for a set of general spacetime alternatives. Section II.3 defines a particular simple spacetime region and solves, for that region, the differential equation involved in the decoherence functional. Section II.4 describes how the effect of that solution depends upon the spacelike surfaces on which the initial and final states are attached. Section II.5 describes a set of three alternatives defined with respect to the region defined in Sec. II.3. Sec. II.6 discusses certain initial and final conditions and their implications, while Section II.7 calculates the decoherence functional for the spacetime coarse graining defined in Sec. II.5. For one class of initial state, the full set of alternatives decoheres and allows prediction of spacetime probabilities, while for a wider class of initial state, decoherence can be achieved by a coarser graining which combines two of the three alternatives into one. Section II.8 summarizes these results and contemplates their shortcomings.

II.2 Decoherence functional and class operators

II.2.1 General prescription

To construct a generalized quantum mechanics of a free relativisitic particle, I follow the approach of [24] in applying the sum-over-histories prescription of Sec. I.3.3 to the parametrized theory described in Sec. I.4.6. The decoherence functional has the form (I.3.12–III.2.17), with the configuration space variables q of the theory realized as spacetime points x. Although we do not presuppose the existence of a Hilbert space of wave functions (in fact the inner product we define below will not be positive definite, it is illustrative to think of the conditions as being described by initial and final "density matrices" as defined in (I.3.16).

We express the action for a free relativistic particle in $(D{+}1)$-dimensional Minkowski spacetime in the canonical form

$$S_{\text{can}} = \int_0^1 d\lambda \left(p \cdot \frac{dx}{d\lambda} - N\,\frac{p^2+m^2}{2m} \right), \tag{I.4.16}$$

where $p^2 = p \cdot p = p^\mu p_\mu = -(p^0)^2 + \mathbf{p}^2$; the fine-grained histories summed over are parametrized paths $\{p(\lambda), x(\lambda)\}$ through phase space and multiplier histories $N(\lambda)$. The multiplier N is a quantity which classically (i.e., for the path of least action) defines the relationship between proper time and the arbitrary parameter λ: $N = \frac{d\tau}{d\lambda}$. Note that the paths are allowed to move forward and backward in the "time" coördinate x^0. This set of fine-grained histories is Lorentz invariant, as opposed to a theory which restricts the paths to move forward in time in a given Lorentz frame.

Note also that the action is invariant under reparametrizations of the parameter λ, if N transforms as the derivative of an invariant quantity. Since only reparametrization-invariant coarse grainings are considered as being physically meaningful, one may restrict the sum over histories to those histories which satisfy the "gauge condition" $\frac{dN}{d\lambda} = 0$. In this gauge, one needs only integrate over a single N, which is the total proper time of the path. The theory will turn out to have a closer correspondence to field theory if one integrates only over positive values[1] of N. The class operator is thus defined by

$$\langle x'' \| C_\alpha \| x' \rangle = \int_0^\infty dN \int\limits_{x''\alpha x'} \mathcal{D}^{D+1}x\,\mathcal{D}^{D+1}p \exp\left[i \int_0^N d\tau \left(p \cdot \frac{dx}{d\tau} - \frac{p^2+m^2}{2m} \right) \right]. \tag{II.2.1}$$

[We only wish to consider coarse grainings which restrict the configuration space path $x(\lambda)$, but it is useful to express the sum over histories in terms of phase space histories because the measure for the path integral is then naturally defined.]

[1] See [31] for some consequences of allowing N to take positive and negative values.

To specify the inner product $\circ$ we define an "initial" spacelike D-surface σ' and a "final" spacelike D-surface σ'' to the future of the initial one, and apply the Klein-Gordon inner product on those surfaces:

$$\Phi(x') \circ \Psi(x') = \int_{\sigma'} d^D\Sigma'^\nu \Phi^*(x') i \overleftrightarrow{\nabla}'_\nu \Psi(x') \tag{II.2.2a}$$

and

$$\Phi(x'') \circ \Psi(x'') = \int_{\sigma''} d^D\Sigma''^\mu \Phi^*(x'') i \overleftrightarrow{\nabla}''_\mu \Psi(x''). \tag{II.2.2b}$$

(Here $\overleftrightarrow{\nabla}$ is the usual bidirectional derivative: $\Phi \overleftrightarrow{\nabla}_\mu \Psi = \Phi \nabla_\mu \Psi - \Psi \nabla_\mu \Phi$) Thus[2]

$$\langle \Phi_i | C_\alpha | \Psi_j \rangle = \int_{\sigma''} d^D\Sigma''^\mu \int_{\sigma'} d^D\Sigma'^\nu \Phi_i^*(x'') i \overleftrightarrow{\nabla}''_\mu \langle x'' \| C_\alpha \| x' \rangle i \overleftrightarrow{\nabla}'_\nu \Psi_j(x'). \tag{II.2.3}$$

Integrating over all paths gives the unrestricted propagator

$$\langle x'' \| C_u \| x' \rangle = 2mi\Delta_F(x''{-}x'), \tag{II.2.4}$$

where

$$\Delta_F(x''{-}x') = \int \frac{d^{D+1}p}{(2\pi)^{D+1}} \frac{e^{ip\cdot(x''-x')}}{-(p^2+m^2)+i\varepsilon} \tag{II.2.5}$$

is the Feynman propagator, which propagates positive energy solutions forward in time and annihilates negative energy solutions:

$$\langle x'' \| C_u \| x' \rangle \circ e^{-i\omega_{\mathbf{p}} t'} e^{i\mathbf{p}\cdot\mathbf{x}'} = 2m e^{-i\omega_{\mathbf{p}} t''} e^{i\mathbf{p}\cdot\mathbf{x}''} \tag{II.2.6a}$$

$$\langle x'' \| C_u \| x' \rangle \circ e^{i\omega_{\mathbf{p}} t'} e^{i\mathbf{p}\cdot\mathbf{x}'} = 0 \tag{II.2.6b}$$

assuming $t'' > t'$ (where $\omega_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2}$). The restriction of the multiplier N to positive values has given the advertized correspondence to field theory, as our propagator is the familiar Feynman propagator. This has also led to the bias towards positive energy solutions (II.2.6).

II.2.2 Spacetime alternatives

As an example of a simple spacetime coarse graining, we define a spacetime region S, and a set of two exclusive and exhaustive alternatives as follows: c_s is the class of paths which at some point enter S, and $c_{\overline{s}}$ is the class of paths which never enter it. (See Fig. II.1.) If we define

$$\langle x'' N \| C_{\overline{s}} \| x' 0 \rangle = \int_{x''\overline{s}x'} \mathcal{D}^{D+1}x\, \mathcal{D}^{D+1}p \exp\left[i \int_0^N d\tau \left(p \cdot \frac{dx}{d\tau} - \frac{p^2+m^2}{2m} \right) \right], \tag{II.2.7}$$

[2] We have, of course, treated the class operator $\langle x'' \| C_\alpha \| x' \rangle$ as a "matrix" and not taken its complex conjugate to apply the inner product $\circ$.

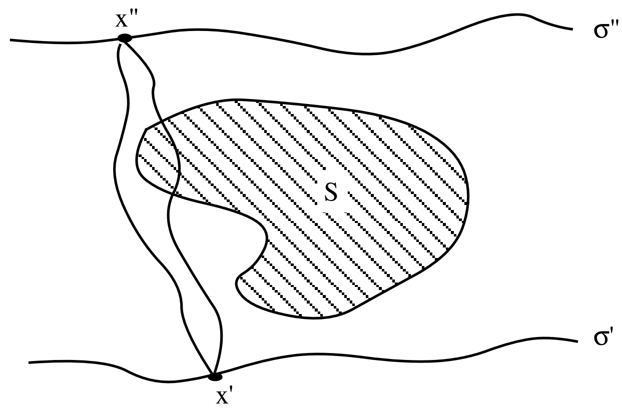


Figure II.1: An example of a spacetime coarse graining. The path on the left never enters the spacetime region S and is thus in the class $c_{\overline{s}}$. The path on the right spends part of its trajectory in S and is thus in the class c_s. ($D-1$ of the D space dimensions have been suppressed.)

so that

$$\langle x''\|C_{\overline{s}}\|x'\rangle = \int_0^\infty dN \langle x''N\|C_{\overline{s}}\|x'0\rangle, \tag{II.2.8}$$

comparing (II.2.7) to the path integral expression for a nonrelativistic propagator, one can show (see [24] for more details) that $\langle x''N\|C_{\overline{s}}\|x'0\rangle$ obeys a five-dimensional Schrödinger-like equation

$$\left(-i\frac{\partial}{\partial N} + \frac{-\nabla^2_{x''}+m^2}{2m} - iE_S(x'')\right)\langle x''N\|C_{\overline{s}}\|x'0\rangle = 0 \tag{II.2.9a}$$

with initial condition

$$\langle x''0\|C_{\overline{s}}\|x'0\rangle = \delta^{D+1}(x''-x')e^{-E_S(x')}, \tag{II.2.9b}$$

where we explicitly allow for the possibility that the region S intersects the initial slice σ' or the final slice σ''. Here

$$E_S(x) = \begin{cases} 0, & x \notin S \\ \infty, & x \in S \end{cases} \tag{II.2.10}$$

is the excluding potential for the region S. Note that

$$e^{-E_S(x)} = \begin{cases} 1, & x \notin S \\ 0, & x \in S. \end{cases} \tag{II.2.11}$$

Equation (II.2.9) is equivalent to the homogeneous PDE

$$\left(-i\frac{\partial}{\partial N} + \frac{-\nabla^2_{x''}+m^2}{2m}\right)\langle x''N\|C_{\overline{s}}\|x'0\rangle = 0, \quad x'' \notin S \tag{II.2.12a}$$

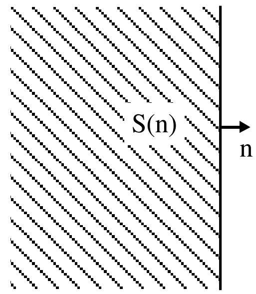


Figure II.2: The region $S(n)$ defined by the unit vector n. ($D-1$ of the D space dimensions have been suppressed.)

with boundary condition

$$\langle x''N\|C_{\overline{s}}\|x'0\rangle = 0, \quad x'' \in \partial S \tag{II.2.12b}$$

and initial condition

$$\langle x''0\|C_{\overline{s}}\|x'0\rangle = \delta^{D+1}(x''-x')e^{-E_S(x')}. \tag{II.2.12c}$$

II.3 Solution by method of images

For a sufficiently simple region, we can construct the class operator $C_{\overline{s}}$ by the method of images. Let n be a constant spacelike unit vector ($n \cdot n = 1$), and $x_n = n \cdot x$ be the component of x along n. Then define $S(n)$ by $x_n \leq 0$ (Fig. II.2), so that $e^{-E_{S(n)}(x)} = \Theta(x_n)$. If we define[3] the reflection of x through the plane $x_n = 0$ by $x_c = x - 2x_n n$, $\langle x''N\|C_u\|x'0\rangle - \langle x''N\|C_u\|x'_c0\rangle$ satisfies (II.2.12a) (by the principle of superposition) and (II.2.12b), and has initial value

$$\langle x''0\|C_u\|x'0\rangle - \langle x''0\|C_u\|x'_c0\rangle = \delta^{D+1}(x''-x') - \delta^{D+1}(x''-x'_c), \tag{II.3.1}$$

which is equal to $\delta^{D+1}(x''-x')$ for $x', x'' \notin S(n)$. Thus

$$\langle x''N\|C_{\overline{s}(n)}\|x'0\rangle = \Theta(x'_n)\Theta(x''_n)\left(\langle x''N\|C_u\|x'0\rangle - \langle x''N\|C_u\|x'_c0\rangle\right) \tag{II.3.2}$$

solves (II.2.12), and yields the class operator

$$\langle x''\|C_{\overline{s}(n)}\|x'\rangle = 2mi\Theta(x'_n)\Theta(x''_n)\left[\Delta_F(x''-x') - \Delta_F(x''-x'_c)\right]. \tag{II.3.3}$$

II.4 Dependence on initial and final time slices

Since our construction (II.2.3) of the matrix elements $\{\langle\Phi_i|C_\alpha|\Psi_j\rangle\}$ from the class operator $\langle x''\|C_\alpha\|x'\rangle$ makes explicit reference to a choice of nonintersecting spacelike surfaces

[3] To avoid confusion, keep in mind that x_n is just a number, while x_c is a $(D+1)$-vector.

σ' and σ'', those matrix elements and hence the decoherence functional could, in principle, depend on the choice of surfaces, and we would like to determine what, if any, that dependence is. Observe that for a given surface σ with normal vector u, the Klein-Gordon inner product (II.2.2) on that surface depends only on the values on σ of the wave function Ψ and its first normal derivative $u^\mu\nabla_\mu\Psi$. Thus the construction of the decoherence functional (I.3.12) depends only on the values on σ'' of $\Phi_i(x'')$ and $u''^\mu\nabla''_\mu\Phi_i(x'')$ and the values on σ' of $\Psi_j(x')$ and $u'^\nu\nabla'_\nu\Psi_j(x')$. To discuss the behavior of the decoherence functional under changes of σ' or σ'', we need to define how the wave functions Φ and Ψ vary off of those surfaces, and we do so by requiring them to satisfy the Klein-Gordon equation. If Ψ and Φ are viewed as functions over all spacetime, this can be seen as the operator version of the mass shell constraint (I.4.15).

Now we can consider how $\langle\Phi_i|C_{\overline{s}}\|x'\rangle = \Phi_i(x'')\circ\langle x''\|C_{\overline{s}}\|x'\rangle$ varies under changes of σ''. As a consequence of (II.2.9a) the class operator $\langle x''\|C_{\overline{s}}\|x'\rangle$ will satisfy the following (for any region S):

$$\left(\frac{-\nabla^2_{x''}+m^2}{2m}\right)\langle x''\|C_{\overline{s}}\|x'\rangle = 0, \quad x'\neq x''\notin S \tag{II.4.1a}$$

$$\langle x''\|C_{\overline{s}}\|x'\rangle = 0, \quad x''\in S. \tag{II.4.1b}$$

We assume here, as throughout this work, that the surfaces σ' and σ'' do not intersect one another, so that $x'\neq x''$ holds as far as we are concerned. Thus $\langle x''\|C_{\overline{s}}\|x'\rangle$ satisfies the Klein-Gordon equation on x'' everywhere except on the boundary ∂S. Since the final wave functions $\{\Phi_i\}$ are taken to be solutions to the Klein-Gordon equation, the usual demonstration of invariance of the Klein-Gordon inner product tells us that we can deform the surface σ'' without changing $\langle\Phi_i|C_{\overline{s}}\|x'\rangle$ so long as its intersection $\sigma''\cap\partial S$ with the boundary of S stays fixed. Examining the behavior of the sum-over-histories construction (II.2.1) under the substitutions $\upsilon = N-\tau$, $y(\upsilon) = x(N-\upsilon)$ and $k=-p$, we see that the class operator is symmetric under the interchange of ends of the path ($\langle x''\|C_\alpha\|x'\rangle = \langle x'\|C_\alpha\|x''\rangle$) so long as the class c_α does not distinguish one end of the path from the other. The class $c_{\overline{s}}$ is such a class.[4] Thus $\langle x''\|C_{\overline{s}}\|x'\rangle$ must satisfy the analogous properties to (II.4.1) with respect to the other argument x'. Thus changes of σ' which leave $\sigma'\cap\partial S$ unchanged will not change $\langle x''\|C_{\overline{s}}|\Psi_j\rangle = \langle x''\|C_{\overline{s}}\|x'\rangle\circ\Psi_j(x')$ either. Since $\langle x''\|C_s\|x'\rangle + \langle x''\|C_{\overline{s}}\|x'\rangle = \langle x''\|C_u\|x'\rangle = 2mi\Delta_F(x''-x')$ by (I.3.14b), and the Feynman propagator satisfies the Klein-Gordon equation on its (nonvanishing) argument, $\langle x''\|C_s\|x'\rangle$ will satisfy the equation whenever $\langle x''\|C_{\overline{s}}\|x'\rangle$ does, and all elements of the decoherence functional will be unchanged under any change of σ' and σ'' which leaves their intersection with ∂S unchanged. (Fig. II.3)

This argument has previously been used [24] to show that the decoherence functional is independent of the choice of nonintersecting surfaces so long as σ' lies completely to the past and σ'' completely to the future of S. The nature of the region $S(n)$ defined in Sec. II.3 prevents us from choosing initial and final spacelike surfaces which do not intersect $S(n)$. What we can do without changing the decoherence functional is generate the D-surface σ from the $(D-1)$-surface $\sigma\cap\partial S(n)$ via curves everywhere tangent to n. (Fig. II.4) Then n

[4] An example of a class which *does* distinguish one end of the class from the other is one which refers to the first time in its trajectory that a particle crosses a surface or enters a region.

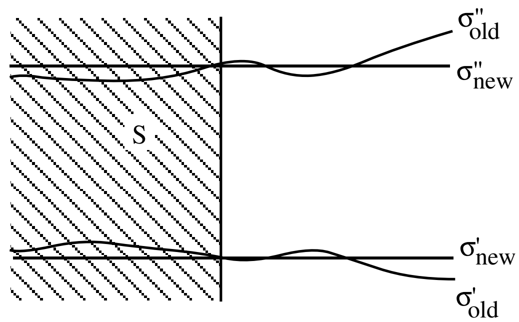


Figure II.3: Varying the surfaces σ' and σ'' on which the inner product (II.2.2) is imposed does not change the decoherence functional, as long as their intersections with ∂S are unchanged. ($D-1$ of the D space dimensions have been suppressed.)

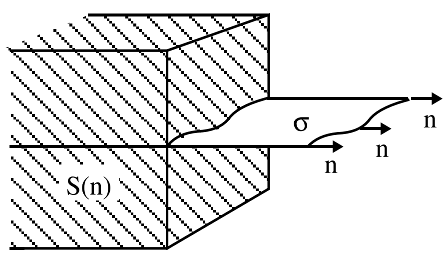


Figure II.4: Generating the surface σ from its intersection with $\partial S(n)$ by projecting along n. ($D-2$ of the D space dimensions have been suppressed.) If $D=1$, $\sigma \cap \partial S(n)$ is a point and σ generated in this fashion will always be flat. With two or more space dimensions, σ will only be flat if $\sigma \cap \partial S(n)$ is; if $\sigma \cap \partial S(n)$ is "wavy", σ will be translationally invariant along n, resembling a sheet of corrugated metal.

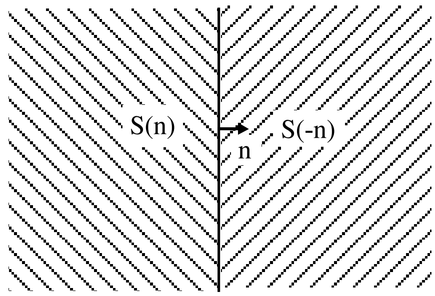


Figure II.5: The regions $S(n)$ ("left") and $S(-n)$ ("right") defined by the unit vector n, along with their common boundary, the "wall" $x_n = 0$. ($D-1$ of the D space dimensions have been suppressed.)

will lie in the surface at all points, and $n^\mu d^D\Sigma_\mu = 0$. [5] This will later prove crucial.

II.5 Our chosen set of alternatives

We can take advantage of the fact that for a given normal vector n, the regions $S(n)$ ($n \cdot x \leq 0$) and $S(-n)$ ($-n \cdot x \leq 0$) are on opposite sides of the same boundary $x_n = 0$. (Fig. II.5) Loosely calling $S(n)$ the "left" side and $S(-n)$ the "right" side of the "wall" $x_n = 0$, we can define a set of alternatives by the answers to the two questions "does the particle ever enter $S(n)$ ($x_n \leq 0$)?" and "does the particle ever enter $S(-n)$ ($x_n \geq 0$)?" The class $c_{\overline{s}(n)} \cap c_{\overline{s}(-n)}$, corresponding to both answers being "no", is empty. The three nontrivial alternatives are: $c_\ell = c_{s(n)} \cap c_{\overline{s}(-n)} = c_{\overline{s}(-n)}$, in which the particle is on the left side of the wall throughout its entire trajectory; $c_r = c_{\overline{s}(n)} \cap c_{s(-n)} = c_{\overline{s}(n)}$, in which the particle is always on the right side; and $c_b = c_{s(n)} \cap c_{s(-n)}$, in which the particle spends some time on each side of the wall, and crosses it in between. This set of three alternatives, illustrated in Fig. II.6, is exhaustive and mutually exclusive, and is thus a suitable coarse graining. The class operators for c_ℓ and c_r were calculated in Sec. II.3, and are given by

$$\langle x''\|C_\ell\|x'\rangle = \langle x''\|C_{\overline{s}(-n)}\|x'\rangle = 2mi\Theta(-x'_n)\Theta(-x''_n)\left[\Delta_F(x''-x') - \Delta_F(x''-x'_c)\right] \quad \text{(II.5.1a)}$$
$$\langle x''\|C_r\|x'\rangle = \langle x''\|C_{\overline{s}(n)}\|x'\rangle = 2mi\Theta(x'_n)\Theta(x''_n)\left[\Delta_F(x''-x') - \Delta_F(x''-x'_c)\right], \quad \text{(II.5.1b)}$$

where we have used the fact that $x_{-n} = -n \cdot x = -x_n$ [and also that x_c is defined the same way with respect to n and $-n$: $x_c = x - 2nx_n = x + 2nx_{-n} = x - 2(-n)x_{-n}$]. The class

[5] Note that in $1+1$ dimensions, this allows us to choose our surface to be a surface of constant time in the reference frame where $n^0 = 0$.

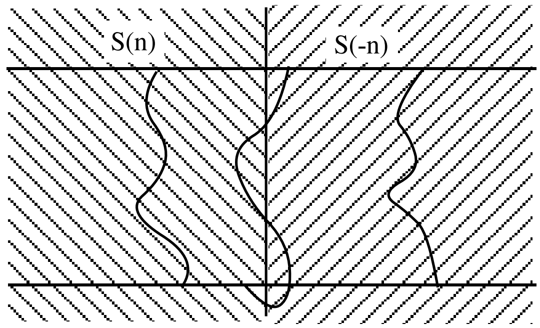


Figure II.6: The coarse graining described in Sec. II.5. The three paths shown are representatives of, from left to right: the class c_ℓ of paths which lie completely to the left of the wall; the class c_b of paths which spend some time on each side of the wall; and the class c_r of paths which lie completely to the right of the wall. ($D-1$ of the D space dimensions have been suppressed.) Compare Fig. 3 of [1].

operator for c_b can be calculated from the superposition law (I.3.14b):

$$\begin{aligned}\langle x''\|C_b\|x'\rangle &= \langle x''\|C_u\|x'\rangle - \langle x''\|C_\ell\|x'\rangle - \langle x''\|C_r\|x'\rangle \\ &= 2mi\{[\Theta(x'_n)\Theta(-x''_n) + \Theta(-x'_n)\Theta(x''_n)]\Delta_F(x''-x') \\ &\qquad + [\Theta(x'_n)\Theta(x''_n) + \Theta(-x'_n)\Theta(-x''_n)]\Delta_F(x''-x'_c)\}. \end{aligned} \tag{II.5.1c}$$

II.6 Properties for certain initial and final conditions

II.6.1 Pure initial state

If we specialize to a pure initial state $\Psi(x')$, it becomes useful to define the branch wave function

$$\Psi_\alpha(x'') = \frac{1}{2m}\langle x''\|C_\alpha|\Psi\rangle = \frac{1}{2m}\langle x''\|C_\alpha\|x'\rangle \circ \Psi(x'), \tag{II.6.1}$$

so that the decoherence functional (I.3.12) has elements

$$D(\alpha, \alpha') = \frac{\Psi_{\alpha'} \circ \rho'' \circ \Psi_\alpha}{\Psi^+ \circ \rho'' \circ \Psi^+}. \tag{II.6.2}$$

Here Ψ^+ is the positive energy part of Ψ [see (II.2.6)]:

$$\Psi^+(x'') = i\Delta_F(x''-x') \circ \Psi(x') = \frac{1}{2m}\langle x''\|C_u|\Psi\rangle, \tag{II.6.3}$$

and is the branch wave function corresponding to the class c_u of all paths. The superposition property (I.3.14) for class operators and the definition (II.6.1) of the branch wave function imply an analogous superposition law for branch wave functions:

$$\Psi_{\overline{\alpha}}(x'') = \sum_{\alpha\in\overline{\alpha}} \Psi_\alpha(x'') \tag{II.6.4a}$$

$$\sum_\alpha \Psi_\alpha(x'') = \Psi^+(x''). \tag{II.6.4b}$$

We postpone for the moment discussion of the final condition ρ''.

The branch wave functions for the classes c_ℓ, c_r and c_b can be given in terms of the branch wave functions $\Psi_{\overline{s}(\pm n)}$ by

$$\Psi_\ell(x'') = \Psi_{\overline{s}(-n)}(x'') \tag{II.6.5a}$$

$$\Psi_r(x'') = \Psi_{\overline{s}(n)}(x'') \tag{II.6.5b}$$

$$\Psi_b(x'') = \Psi^+(x'') - \Psi_\ell(x'') - \Psi_r(x''). \tag{II.6.5c}$$

Using (II.3.3), we write $\Psi_{\overline{s}(\pm n)}(x'')$ as

$$\Psi_{\overline{s}(\pm n)}(x'') = \Theta(\pm x''_n)\int_{\sigma'} d^D\Sigma'^\nu \Theta(\pm x'_n)\left[i\Delta_F(x''-x') - i\Delta_F(x''-x'_c)\right] i\overleftrightarrow{\nabla}'_\nu \Psi(x'). \tag{II.6.6}$$

As described in Sec. II.4, we can, without loss of generality, choose σ' to satisfy $n_\nu d^D\Sigma'^\nu = 0$, which allows us to move the $\Theta(\pm x'_n)$ to the other side of the $\overleftrightarrow{\nabla}'_\nu$ [since $\nabla_\nu\Theta(\pm x'_n) = \pm n_\nu \delta(x'_n)$, which is orthogonal to $d^D\Sigma'^\nu$] and get

$$\Psi_{\overline{s}(\pm n)}(x'') = \Theta(\pm x''_n)\int_{\sigma'} d^D\Sigma'^\nu \left[i\Delta_F(x''-x') - i\Delta_F(x''-x'_c)\right] i\overleftrightarrow{\nabla}'_\nu \Theta(\pm x'_n)\Psi(x'). \tag{II.6.7}$$

If we change the integration variable from x' to x'_c in the second term of the integral (which we can do because the construction of σ' ensures that $x'_c \in \sigma'$ if and only if $x' \in \sigma'$), we obtain

$$\Psi_{\overline{s}(\pm n)}(x'') = \Theta(\pm x''_n)\int_{\sigma'} d^D\Sigma'^\nu i\Delta_F(x''-x') i\overleftrightarrow{\nabla}'_\nu \left[\Psi(x')\Theta(\pm x'_n) - \Psi(x'_c)\Theta(\mp x'_n)\right] \tag{II.6.8}$$

Without an additional restriction on $\Psi(x')$, it is quite difficult to proceed any further.

Antisymmetric initial state

If we choose our initial state to be an odd function of x_n (which we write as Υ to distinguish it from the generic initial state Ψ):

$$\Upsilon(x_c) = -\Upsilon(x), \tag{II.6.9}$$

we have $\Upsilon(x')\Theta(\pm x'_n) - \Upsilon(x'_c)\Theta(\mp x'_n) = \Upsilon(x')$, and (II.6.8) becomes

$$\Upsilon_{\overline{s}(\pm n)}(x'') = \Theta(\pm x''_n)\int_{\sigma'} d^D\Sigma'^\nu i\Delta_F(x''-x') i\overleftrightarrow{\nabla}'_\nu \Upsilon(x') = \Theta(\pm x''_n)\Upsilon^+(x''). \tag{II.6.10}$$

Thus the branch wave functions for this initial state are

$$\Upsilon_\ell(x'') = \Upsilon_{\overline{s}(-n)}(x'') = \Theta(-x''_n)\Upsilon^+(x'') \tag{II.6.11a}$$
$$\Upsilon_r(x'') = \Upsilon_{\overline{s}(n)}(x'') = \Theta(x''_n)\Upsilon^+(x'') \tag{II.6.11b}$$
$$\Upsilon_b(x'') = 0. \tag{II.6.11c}$$

Note that we can construct a Klein-Gordon state satisfying the antisymmetry property (II.6.9) throughout all spacetime by taking any Klein-Gordon state $\Omega(x)$ which is not symmetric about $x_n = 0$ and defining $\Upsilon(x) = \frac{1}{2}\left[\Omega(x) - \Omega(x_c)\right]$, and note also that both the positive and negative energy parts of Υ have the antisymmetry property as well.

Initial state with restricted support

Another technique for simplifying the branch wave functions, used on the nonrelativistic particle by Yamada and Takagi [2] is to choose an initial state which vanishes either in or out of the region S. Since we attach the initial state with the Klein-Gordon inner product, we need to go a step further, and require that both the initial state $\Psi(x')$ and its normal derivative $u'^\nu \nabla'_\nu \Psi(x')$ vanish on the appropriate part of the initial surface. For brevity's sake, we define the "support" of a wave function to be anywhere where the wave function or its normal derivative is nonvanishing. Thus we want to construct a wave function whose support on the initial surface σ' is confined to (say) the left side of the wall ($x_n < 0$). It is always possible to construct a solution to the Klein-Gordon equation $\Psi(x)$ which has an arbitrary value $f(x')$ and normal derivative $g(x')$ on a surface σ', but it will in general be necessary to construct it out of both positive and negative energy components.[6]

If we construct an initial state (which we call Ξ) whose support on the surface σ' is confined to the left side of the wall:

$$\Xi(x') = 0 = u' \cdot \nabla' \Xi(x') \text{ when } x' \in \sigma \text{ and } x'_n \geq 0 \tag{II.6.12}$$

(see Fig. II.7), then $\Theta(x'_n)\Xi(x')$ and its normal derivative vanish and (II.6.7) gives

$$\Xi_r(x'') = \Xi_{\overline{s}(n)}(x'') = 0. \tag{II.6.13a}$$

Turning the tables and considering the effect the semi-infinite support property (II.6.12) has on $\Xi_\ell = \Xi_{\overline{s}(-n)}$, we see that $\Theta(-x'_n)\Xi(x')$ has the same value and normal derivative on σ' as Ξ itself, and we will be able to drop the $\Theta(-x'_n)$ from (II.6.7), and obtain

$$\Xi_\ell(x'') = \Xi_{\overline{s}(-n)}(x'') = \Theta(-x''_n)\left[\Xi^+(x'') - \Xi^+(x''_c)\right]. \tag{II.6.13b}$$

[We have used the easily proved result that $\Delta_F(x''-x'_c) = \Delta_F(x''_c-x')$.]

The wavefunction Ξ_b can again be found by superposition, and is given by:

$$\Xi_b(x'') = \Theta(x''_n)\Xi^+(x'') + \Theta(-x''_n)\Xi^+(x''_c). \tag{II.6.13c}$$

[6] I am indebted to R. S. Tate for pointing this out to me.

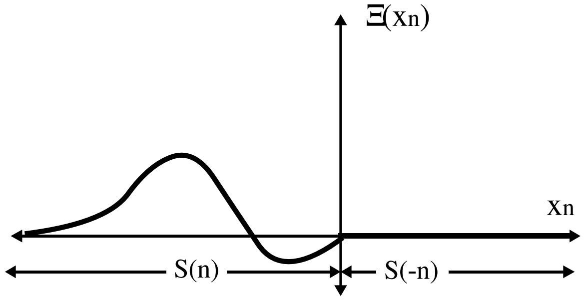


Figure II.7: Schematic plot of a wave function Ξ whose support on σ' is confined to $x_n < 0$. This is a plot of Ξ as a function of x_n for fixed $\mathbf{x}_\perp$ on the surface σ'. Note that $u'^{\nu}\nabla'_{\nu}\Xi(x')$ must also vanish on the "right" half of the surface σ for Ξ to have semi-infinite support as defined in Sec. II.6.1. [See (II.6.12).]

II.6.2 Future indifference

In order to evaluate the decoherence functional (II.6.2) we need to consider the final condition ρ''. In analogy with our observations that the universe has a preferred time direction, we would like to abandon the time-symmetric construction of (I.3.12) and choose a condition of future indifference, i.e., a completely unspecified final condition. In most time-symmetric formulations of quantum mechanics, this condition is implemented by replacing the final density matrix with the identity operator, so that $\Psi_{\alpha'} \circ \rho'' \circ \Psi_\alpha \to \Psi_{\alpha'} \circ \Psi_\alpha$, but this cannot be the prescription here, since it is not manifestly positive when $\alpha = \alpha'$, as our initial construction was.

To see why this fails, construct completely unspecified density matrices for the positive and negative energy sectors of the theory:

$$\rho_\pm(x_2, x_1) = \int \frac{d^D p}{(2\pi)^D 2\omega_{\mathbf{p}}} e^{\mp i\omega_{\mathbf{p}}(t_2 - t_1)} e^{i\mathbf{p}\cdot(\mathbf{x}_2 - \mathbf{x}_1)}. \tag{II.6.14}$$

They have the following property under the Klein-Gordon inner product:

$$\rho_\pm(x_2, x_1) \circ \Psi(x_1) = \pm\Psi^\pm(x_2), \tag{II.6.15}$$

where $\Psi(x)$ is any solution to the Klein-Gordon equation, and $\Psi^+(x)$ and $\Psi^-(x)$ are its positive and negative energy components, respectively $[\Psi(x) = \Psi^+(x) + \Psi^-(x)]$. The "identity operator" with respect to this inner product is thus $\rho_+ - \rho_-$. It is unsuitable for a final condition ρ'', since some of the weights $\{p''_i\}$ it implies are negative, in violation of the rules set out in Sec. I.3.3. Instead, we take our condition of future indifference to be

$$\rho_{\rm fi} = \rho_+ + \rho_-, \tag{II.6.16}$$

so that[7]

$$\Psi_{\alpha'} \circ \rho_{\rm fi} \circ \Psi_\alpha = \Psi^+_{\alpha'} \circ \Psi^+_\alpha - \Psi^-_{\alpha'} \circ \Psi^-_\alpha. \tag{II.6.17}$$

[7]Technically speaking, we should not talk about the positive and negative energy components of the

This is equivalent to the result we would have gotten if we had used the positive definite inner product for Klein-Gordon wave functions, and then chosen the identity as our final density matrix. This inner product is nonlocal in the spacetime coördinate x, so, for example, wave functions which do not overlap can still have a nonvanishing inner product.

Note that we can replace the normalization factor $\Psi^+ \circ \rho'' \circ \Psi^+$ in (II.6.2) with $\Psi^+ \circ \Psi^+$ if we use the final condition (II.6.16). It will therefore prove useful to normalize our initial wave function so that

$$\Psi^+ \circ \Psi^+ = 1. \tag{II.6.18}$$

The decoherence functional is then

$$D(\alpha, \alpha') = \Psi_{\alpha'} \circ \rho_{\text{fi}} \circ \Psi_\alpha. \tag{II.6.19}$$

II.7 Results

II.7.1 Results for antisymmetric initial state

Using the antisymmetric initial state Υ from Sec. II.6.1, the branch wave functions for the three classes are

$$\Upsilon_\ell(x'') = \Theta(-x''_n)\Upsilon^+(x'') \tag{II.6.11a}$$
$$\Upsilon_r(x'') = \Theta(x''_n)\Upsilon^+(x'') \tag{II.6.11b}$$
$$\Upsilon_b(x'') = 0. \tag{II.6.11c}$$

The elements of the decoherence functional (II.6.19) are calculated in Appendix II.A, and found (when the final surface σ'' is taken to be one of constant time t'') to be

$$\begin{pmatrix} D(\ell,\ell) = \frac{1}{2} + \Delta D & D(\ell,r) = -\Delta D & D(\ell,b) = 0 \\ D(r,\ell) = -\Delta D & D(r,r) = \frac{1}{2} + \Delta D & D(r,b) = 0 \\ D(b,\ell) = 0 & D(b,r) = 0 & D(b,b) = 0 \end{pmatrix} \tag{II.7.1}$$

where[8]

$$\Delta D = 2\int \frac{dk_{1n}dk_{2n}d^{D-1}k_\perp}{(2\pi)^2}\frac{\omega_1+\omega_2}{2\sqrt{\omega_1\omega_2}}\widetilde{\Upsilon}^+(\mathbf{k}_2)^*\widetilde{\Upsilon}^+(\mathbf{k}_1)\frac{e^{-i(\omega_1-\omega_2)t''}}{k_{1n}-k_{2n}}\ln\left(\frac{\omega_1-k_{1n}}{\omega_2-k_{2n}}\right). \tag{II.7.1a}$$

Aside from $D(\ell, r) = D(r, \ell) = -\Delta D = \Upsilon_\ell \circ \rho_{\text{fi}} \circ \Upsilon_r$, all of the off-diagonal elements vanish (this is true for any final condition, in fact). $D(\ell, r) = D(r, \ell)$ generally does not vanish,

branch wave functions $\{\Psi_\alpha\}$, since we showed in Sec. II.4 that the class operators (and hence the branch wave functions) are guaranteed to satisfy the Klein Gordon equation only when $x'' \notin \partial S$, and the branch wave functions are thus not in the space of solutions to the Klein-Gordon equation. However, a more careful analysis (see Appendix II.A) shows that, defining $\Psi^\pm$ by (II.6.15), $\Phi\circ\Psi = \Phi^+\circ\Psi^+ + \Phi^-\circ\Psi^- = (\Phi^+ + \Phi^-)\circ(\Psi^+ + \Psi^-)$ (where all inner products are taken on the same surface), even if Φ and Ψ are not solutions to the Klein-Gordon equation. The division into positive and negative energy parts is thus well-defined for our purposes.

[8] We use here several pieces of notation defined in Appendix II.A, namely $\mathbf{v}_\perp = \mathbf{v} - v_n\mathbf{n}$ and $\omega_\perp = \sqrt{\mathbf{k}_\perp^2 + m^2}$ (so that $\omega = \sqrt{k_n^2 + \omega_\perp^2}$), and also that $\widetilde{\Upsilon}^+$ is the Fourier transform (II.A.7) of the positive energy part of Υ. We are also working in a reference frame where $\mathbf{n}$ has no time component.

despite the lack of overlap of the branch wave functions, because of the nonlocality of the positive definite inner product induced by the final condition in section II.6.2. Note that whenever the alternatives do decohere ($\Delta D \approx 0$), the probabilities are given by $p(\ell) \approx 1/2 \approx p(r)$, $p(b) = 0$. [Symmetry arguments make it clear that we must have $p(\ell) = p(r)$.] Note also that while the decoherence functional depends on the time t'' of the final surface, it is completely independent of the initial surface σ'.

To determine whether or not we have decoherence, we need to consider further properties of the initial condition Υ (or equivalently its Fourier transform $\widetilde{\Upsilon}$).

Let $\widetilde{\Upsilon}$ be given by a Gaussian wavepacket peaked at $\mathbf{k}_0$, $\mathbf{x}_0$ and t_0, minus its reflection through $k_n = 0$. That is to say

$$\begin{aligned}\widetilde{\Upsilon}(\mathbf{k}) &= Ce^{i\omega_{\mathbf{k}}t_0}\left(e^{-i\mathbf{k}\cdot\mathbf{x}_0}e^{-(\mathbf{k}-\mathbf{k}_0)^2/4(\delta k)^2} - e^{-i\mathbf{k}_c\cdot\mathbf{x}_0}e^{-(\mathbf{k}_c-\mathbf{k}_0)^2/4(\delta k)^2}\right)\\ &= Ce^{-i\mathbf{k}_\perp\cdot\mathbf{x}_0}e^{i\omega_{\mathbf{k}}t_0}e^{-(\mathbf{k}_\perp-\mathbf{k}_{0\perp})^2/4(\delta k)^2}\sum_{\xi=\pm1}\xi e^{-i\xi k_n x_{0n}}e^{-(k_n-\xi k_{0n})^2/4(\delta k)^2},\end{aligned} \tag{II.7.2}$$

where the normalization constant is given by

$$|C|^2 = \frac{1}{2(\delta k\sqrt{2\pi})^D\left[1-e^{-k_{0n}^2/2(\delta k)^2}e^{-x_{0n}^2/2(\delta x)^2}\right]} \tag{II.7.2a}$$

with $\delta x\delta k = 1/2$. We then have

$$\begin{aligned}\Delta D &= |C|^2\int\frac{dk_{1n}dk_{2n}d^{D-1}k_\perp}{(2\pi)^2}\frac{\omega_1+\omega_2}{2\sqrt{\omega_1\omega_2}}e^{-(\mathbf{k}_\perp-\mathbf{k}_{0\perp})^2/2(\delta k)^2}\\ &\qquad\qquad\times e^{-i(\omega_1-\omega_2)(t''-t_0)}\ln\left(\frac{\omega_1-k_{1n}}{\omega_2-k_{2n}}\right)\\ &\quad\times\sum_{\xi_1=\pm1}\sum_{\xi_2=\pm1}\xi_1\xi_2 e^{-(k_{1n}-\xi_1k_{0n})^2/4(\delta k)^2}e^{-(k_{2n}-\xi_2k_{0n})^2/4(\delta k)^2}\frac{e^{-i\xi_1k_{1n}x_{0n}}e^{i\xi_2k_{2n}x_{0n}}}{k_{1n}-k_{2n}}\\ &= 2|C|^2\int\frac{dk_{1n}dk_{2n}d^{D-1}k_\perp}{(2\pi)^2}\frac{\omega_1+\omega_2}{2\sqrt{\omega_1\omega_2}}e^{-(\mathbf{k}_\perp-\mathbf{k}_{0\perp})^2/2(\delta k)^2}e^{-(k_{1n}-k_{0n})^2/4(\delta k)^2}\\ &\quad\times e^{-(k_{2n}-k_{0n})^2/4(\delta k)^2}e^{-i(\omega_1-\omega_2)(t''-t_0)}e^{-i(k_{1n}-k_{2n})x_{0n}}\sum_{\xi=\pm1}\frac{2\xi}{k_{1n}-\xi k_{2n}}\ln\left(\frac{\omega_1-k_{1n}}{\omega_2-\xi k_{2n}}\right),\end{aligned} \tag{II.7.3}$$

where the final form has been arrived at by changing the variables in the integrals $k_{1n} \to \xi_1 k_{1n}$, $k_{2n} \to \xi_2 k_{2n}$ and then making the substitution $\xi = \xi_1\xi_2$.

In the limit that $\delta k \to 0$, we can replace k_{1n} and k_{2n} with k_{0n} and $\mathbf{k}_\perp$ with $\mathbf{k}_{0\perp}$ everywhere except in the Gaussian factors and perform the integrals. We can do this because

$$\lim_{k_{1n}\to k_0}\frac{1}{k_{1n}-k_{0n}}\ln\left(\frac{\omega_1-k_{1n}}{\omega_0-k_{0n}}\right) = -\frac{1}{\omega_0} \tag{II.7.4}$$

is finite, and we obtain

$$\begin{aligned}\Delta D &= \frac{2\,|C|^2}{(2\pi)^2}(\delta k\sqrt{4\pi})^2(\delta k\sqrt{2\pi})^{D-1}\,2\left[-\frac{1}{\omega_0}+\frac{1}{k_{0n}}\ln\left(\frac{\omega_0-k_{0n}}{\omega_{0\perp}}\right)\right]+\mathcal{O}([\delta k]^2)\\ &= -4\frac{\delta k}{(2\pi)^{3/2}}\left[\frac{1}{\omega_0}-\frac{1}{k_{0n}}\ln\left(\frac{\omega_0-k_{0n}}{\omega_{0\perp}}\right)\right]+\mathcal{O}([\delta k]^2).\end{aligned} \tag{II.7.5}$$

Thus we have approximate decoherence to lowest order in δk. Note that the first order correction to the decoherence functional is independent of the time t'' of the final surface.

For a generic antisymmetric initial condition Υ, (II.7.1a) has no reason to be small, so the current set of alternatives will probably not decohere. However, consider a coarser graining in which c_ℓ and c_r are combined into a single class c_o, consisting of all paths which stay on one side or the other of the wall, and never cross it. We can use the superposition property (I.3.4d) to construct the decoherence functional from the finer-grained one (II.7.1). $D(o,o)=D(\ell,\ell)+D(\ell,r)+D(r,\ell)+D(r,r)=1$, etc.

The elements of the decoherence functional are given by

$$\begin{pmatrix} D(o,o)=1 & D(o,b)=0\\ D(b,o)=0 & D(b,b)=0\end{pmatrix} \tag{II.7.6}$$

so we have exact decoherence, and probabilities of 1 for c_o and 0 for c_b. This corresponds to the definite prediction that for a pure initial state antisymmetric about $x_n=0$, the particle path will not cross that surface. Since the antisymmetry property holds throughout all spacetime, this result is independent of the choice of initial and final surfaces.

This last result can be seen from another point of view, allowing a slight generalization. Using the superposition property for branch wave functions (II.6.4), we can construct

$$\Upsilon_o(x'')=\Upsilon_\ell(x'')+\Upsilon_r(x'')=\Upsilon^+(x''). \tag{II.7.7}$$

Recalling that

$$\Upsilon_b(x'')=0, \tag{II.6.11c}$$

we see that all branch wave functions but one vanish. Examination of (II.6.2) shows that whenever this is the case, the only nonvanishing element of the decoherence functional will be the diagonal one corresponding to the alternative with the nonvanishing branch wave function, and we will have decoherence, and a definite prediction of that alternative. This will hold for any final condition [except of course for pathological cases when the final condition is inconsistent with the initial condition ($\Psi\circ\rho''\circ\Psi=0$), in which case the denominator of (II.6.2) vanishes, and the decoherence functional is ill-defined].

II.7.2 Results for initial state with restricted support

With the initial state Ξ from Sec. II.6.1, which vanishes, along with its normal derivative, on the surface $\sigma'\cap S(-n)$, we find that the branch wave functions for the three classes are

$$\Xi_\ell(x'') = \Theta(-x''_n)[\Xi^+(x'') - \Xi^+(x''_c)] \tag{II.6.13b}$$
$$\Xi_r(x'') = 0 \tag{II.6.13a}$$
$$\Xi_b(x'') = \Theta(x''_n)\Xi^+(x'') + \Theta(-x''_n)\Xi^+(x''_c). \tag{II.6.13c}$$

Now the wave functions Ξ_ℓ and Ξ_b overlap, so we do not expect decoherence, even naïvely, unless we coarse grain so that only one of the branch wave functions is nonvanishing. This amounts to recombining c_ℓ and c_b into $c_{s(n)}$, so that the decoherence functional is

$$\begin{pmatrix} D\big(s(n), s(n)\big) = 1 & D\big(s(n), \overline{s}(n)\big) = 0 \\ D\big(\overline{s}(n), s(n)\big) = 0 & D\big(\overline{s}(n), \overline{s}(n)\big) = 0 \end{pmatrix} \tag{II.7.8}$$

which decoheres, with probabilities of 1 for $c_{s(n)}$ and 0 for $c_{\overline{s}(n)}$. Here we have a definite prediction that the particle will at some point in its trajectory be found in $S(n)$. This result, however, depends very much on the choice of the initial surface σ'.

II.8 Discussion

For our simple coarse graining (see Fig. II.6), we were able to calculate explicit expressions for the class operators $C_{s(\pm n)}$ and $C_{\overline{s}(\pm n)}$, and hence for C_ℓ, C_r and C_b.

To calculate branch wave functions for a pure initial state, we chose the state to satisfy special conditions.

- If the wave function Υ was antisymmetric under reflection through $x_n = 0$, the branch wave function Υ_b vanished, while the nonvanishing branches Υ_ℓ and Υ_r had no overlap. This result held no matter what the initial surface σ'.
- If the wave function Ξ and its first normal derivative vanished on that part of the initial surface σ' which was outside of $S(n)$, the branch wave function Ξ_r vanished, but the other two branches, Ξ_ℓ and Ξ_b, overlapped. This held only for one specific choice of σ'

We could not simply take the inner product of branch wave functions to calculate the decoherence functional, since that would have been tantamount to choosing a non-positive-definite final density matrix. Thus even for the initial state Υ, the alternatives c_ℓ and c_r did not automatically decohere just because the branch wave functions did not overlap. If we restricted the final surface to be flat, we could calculate explicit expressions for the elements of the decoherence functional. For some choices of initial state, the off-diagonal elements were small, but in general they could be appreciable. Whenever the alternatives did decohere, the probability for each was 1/2, which we would have predicted on symmetry grounds.

If we coarser grained either example so that only one branch wave function was nonvanishing, we of course found decoherence and a definite prediction (probability 1) of the other alternative, viz.:

- For the initial condition Υ, if the alternatives were chosen to be c_b and $c_o = c_\ell \cup c_r$, we found decoherence for any nonpathological final condition, with probabilities of 0 and

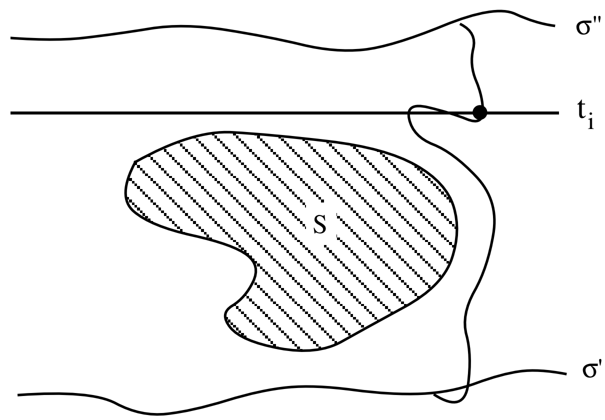


Figure II.8: Dividing up a path which avoids a compact region S. The path from σ' to the last crossing of the intermediate surface t_i is in the class of paths from σ' to t_i which avoid S. The path from the last crossing of t_i to σ'' is in the class of paths from t_i to σ'' which do not cross back over t_i, and can be defined without reference to S. ($D-1$ of the D space dimensions have been suppressed.)

1, respectively. This was a definite prediction that the particle did not cross $x_n = 0$, given an antisymmetric initial condition.

- For the initial condition Ξ, if the alternatives were chosen to be $c_{\overline{s}(n)} = c_r$ and $c_{s(n)} = c_\ell \cup c_b$, we found decoherence for any nonpathological final condition, with probabilities of 0 and 1, respectively. This was a definite prediction that the particle spent part of its trajectory in $S(n)$, given an initial condition which had no support outside of $S(n)$. This is hardly surprising, and it only holds if we attach the initial wave function on the correct hypersurface.

Finally, let us observe that many of our complications were a result of the fact the region which we considered intersected with our initial and final surfaces. If we had considered a region S bounded in time, we could have chosen our initial surface to lie to the past and our final surface to the future of it. As was discussed in Sec. II.4, this would make the decoherence functional necessarily independent of the choice of surface. It would also have eliminated the complications in the choice of the final condition, since the branch wave functions would have been positive energy solutions to the Klein-Gordon equation. The proof is straightforward: construct an intermediate surface of constant time t_i to future of S but the past of σ''. (Section II.4 always allows us to deform the surface σ'' so that such a constant-time surface will "fit" in.) By a construction analogous to that of Halliwell and Ortiz [32], the propagation from σ' to σ'' avoiding the region S can be broken up (at the last crossing of t_i) into propagation from σ' to t_i avoiding S followed by propagation from t_i to σ'' which does not cross back over t_i. (See Fig. II.8.) The class operator can thus be written

$$\langle x''\|C_{\overline{s}}\|x'\rangle = \int d^D x_i \Delta_{1t_i}(x'', x_i)\langle \mathbf{x}_i t_i \|C_s\|x'\rangle \tag{II.8.1}$$

where Δ_{1t_i} is the Newton-Wigner propagator:

$$\Delta_{1t_i}(x, x_i) = \int \frac{d^D p}{(2\pi)^D} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}_i)} e^{-i\omega_{\mathbf{p}}(t-t_i)}. \tag{II.8.2}$$

Since Δ_{1t_i} is constructed from positive-energy solutions of the Klein-Gordon operator, the branch wave functions Ψ_s and $\Psi_{\overline{s}}$ will each be positive-energy solutions themselves. Thus $\Psi_{\alpha'} \circ \rho_- \circ \Psi_\alpha = 0$, so $\Psi_{\alpha'} \circ \rho_{\text{fi}} \circ \Psi_\alpha = \Psi_{\alpha'} \circ \rho_+ \circ \Psi_\alpha = \Psi_{\alpha'} \circ \Psi_\alpha$, and we really do simply calculate the inner product of the branches.

However, it was the simplicity of the region $S(n)$ which allowed us to solve the PDE problem analytically in the first place. Solution of (II.2.12) for finite regions of spacetime cannot be accomplished through straightforward method-of-images or separation-of-variables methods. In the nonrelativistic case, this problem is circumvented for example in the case of a region which extends from t_1 to t_2 by propagating from t' to t_1 with the free propagator, from t_1 to t_2 with the restricted propagator calculated as though the region existed for all time, and then from t_2 to t'' with the free propagator. Since our paths are not single-valued in time, we cannot "turn off" the restricting region before and after we reach it, since we have to include in the sum paths which double back into a previous regime.

II.9 Conclusions

Using the generalized quantum mechanics formalism described by Hartle for the quantum mechanics of the relativistic worldline, we have examined one particularly simple coarse graining. For a suitable choice of initial conditions, albeit a more restrictive one than for the nonrelativistic theory, we were able to assign approximate probabilities to some sets of alternatives.

II.A Appendix: Calculation of $\Upsilon_{\alpha'} \circ \rho_{\mathbf{fi}} \circ \Upsilon_\alpha$

To calculate the elements of the decoherence functional for Sec. II.7.1, we first expand our notational convention for the branches to include $\Upsilon_{-1} \equiv \Upsilon_\ell$ and $\Upsilon_{+1} \equiv \Upsilon_r$ so that we can write $\Upsilon_\lambda(x) = \Theta(\lambda x_n)\Upsilon(x)$, where $\lambda^2 = 1$. The nonvanishing elements of the decoherence functional are now

$$D(\lambda_1, \lambda_2) = \Upsilon_{\lambda_2} \circ \rho_{\text{fi}} \circ \Upsilon_{\lambda_1}, \tag{II.A.1}$$

where the inner product is on the surface σ''.

If Ψ is a solution to the Klein-Gordon equation, we know that $(\rho_+ - \rho_-) \circ \Psi = \Psi_+ + \Psi_- = \Psi$. This will not be true for Υ_λ because it is *not* a solution. However, for the purposes of the Klein-Gordon inner product on the surface σ'', we only need the value and normal derivative of each function on σ''. We can thus replace Υ_λ by Y_λ, a Klein-Gordon wave function[9] which matches Υ_λ and its normal derivative on σ''. This gives us

$$\Upsilon_{\lambda_2} \circ \Upsilon_{\lambda_1} = \Upsilon_{\lambda_2} \circ Y_{\lambda_1} = \Upsilon_{\lambda_2} \circ (\rho_+ - \rho_-) \circ Y_{\lambda_1} = \Upsilon_{\lambda_2} \circ (\rho_+ - \rho_-) \circ \Upsilon_{\lambda_1} \tag{II.A.2}$$

We thus have

$$D(\lambda_1, \lambda_2) = \Upsilon_{\lambda_2} \circ (\rho_+ + \rho_-) \circ \Upsilon_{\lambda_1} = \Upsilon_{\lambda_2} \circ \Upsilon_{\lambda_1} + 2\Upsilon_{\lambda_2} \circ \rho_- \circ \Upsilon_{\lambda_1}. \tag{II.A.3}$$

The first term is simple enough to calculate:

$$\Upsilon_{\lambda_2} \circ \Upsilon_{\lambda_1} = \int_{\sigma''} d^D\Sigma''^\mu \Theta(\lambda_2 x''_n) \Upsilon^+(x'')^* i \overleftrightarrow{\nabla}''_\mu \Theta(\lambda_1 x''_n) \Upsilon^+(x''). \tag{II.A.4}$$

again, since we can choose σ'' to satisfy $n_\mu d^D\Sigma''^\mu = 0$, we can move the step functions through the derivative to get

$$\begin{aligned} \Upsilon_{\lambda_2} \circ \Upsilon_{\lambda_1} &= \int_{\sigma''} d^D\Sigma''^\mu \Theta(\lambda_2 x''_n) \Theta(\lambda_1 x''_n) \Upsilon^+(x'')^* i \overleftrightarrow{\nabla}''_\mu \Upsilon^+(x'') \\ &= \delta_{\lambda_1\lambda_2} \int_{\sigma''} d^D\Sigma''^\mu \Theta(\lambda_1 x''_n) \Upsilon^+(x'')^* i \overleftrightarrow{\nabla}''_\mu \Upsilon^+(x''). \end{aligned} \tag{II.A.5}$$

The symmetry of σ'' and antisymmetry of Υ^+ tell us that $\Upsilon_\ell \circ \Upsilon_\ell = \Upsilon_r \circ \Upsilon_r$, so

$$\Upsilon_{\lambda_2} \circ \Upsilon_{\lambda_1} = \delta_{\lambda_1\lambda_2} \frac{\Upsilon^+ \circ \Upsilon^+}{2} = \frac{\delta_{\lambda_1\lambda_2}}{2}. \tag{II.A.6}$$

To calculate the correction term $\Upsilon_{\lambda_2} \circ \rho_- \circ \Upsilon_{\lambda_1}$, we first define the Fourier transform of Υ^+ by[10]

$$\Upsilon^+(x) = \int \frac{d^Dk}{(2\pi)^{D/2}\sqrt{2\omega_{\mathbf{k}}}} e^{i\mathbf{k}\cdot\mathbf{x}} e^{-i\omega_{\mathbf{k}} t} \widetilde{\Upsilon}^+(\mathbf{k}). \tag{II.A.7}$$

The inner product of two positive energy states is expressed in terms of the Fourier transform by

$$\Phi^+ \circ \Psi^+ = \int d^Dk\, \Phi^+(\mathbf{k})^* \Psi^+(\mathbf{k}), \tag{II.A.8}$$

so the normalization condition (II.6.18) is written as

$$\int d^Dk \left|\Upsilon^+(\mathbf{k})\right|^2 = 1 \tag{II.A.9}$$

In a reference frame where n is has no time component, we can split the spatial part $\mathbf{v}$ of a vector v into components along $\mathbf{n}$: ($v_n = \mathbf{n} \cdot \mathbf{v} = n \cdot v$) and perpendicular to $\mathbf{n}$: ($\mathbf{v}_\perp = \mathbf{v} - v_n \mathbf{n}$). In analogy to v_c defined in Sec. II.3, we define $\mathbf{v}_c = \mathbf{v} - 2v_n\mathbf{n} = -v_n\mathbf{n} + \mathbf{v}_\perp$.

[9] It is straightforward to show that such a wave function exists, and is uniquely given by $Y_\lambda = (\rho_+ - \rho_-) \circ \Upsilon_\lambda$.

[10] Note that the requirement that Υ^+ be a positice energy solution to the Klein-Gordon equation means that the temporal frequency $\omega_{\mathbf{k}}$ of a mode $\widetilde{\Upsilon}(\mathbf{k})$ is determined by its spatial frequency $\mathbf{k}$.

$\widetilde{\Upsilon}^+$ is determined from Υ^+ by

$$\widetilde{\Upsilon}^+(\mathbf{k}) = \sqrt{2\omega_{\mathbf{k}}}e^{i\omega_{\mathbf{k}}t}\int \frac{d^Dx}{(2\pi)^{D/2}}e^{-i\mathbf{k}\cdot\mathbf{x}}\Upsilon^+(x), \tag{II.A.10}$$

so $\widetilde{\Upsilon}$ obeys an antisymmetry property similar to (II.6.9):

$$\widetilde{\Upsilon}^+(\mathbf{k}_c) = -\widetilde{\Upsilon}^+(\mathbf{k}) \tag{II.A.11}$$

To proceed any further, we would like an explicit form for the surface σ''. The simplest would be that σ'' is a surface of constant time t''. However, that condition would not be Lorentz invariant, as it would pick out a reference frame in which the final surface was one of constant time. We know from Sec. II.4 that we are only restricted in the choice of σ'' by the form of the $(D-1)$-surface $\sigma''\cap\partial S(n)$. If we restrict our attention to choices of $\sigma''\cap\partial S(n)$ which are flat (a suitably invariant condition), we can always work in a reference frame in which σ'' *is* a surface of constant time. Since we construct σ'' so that n lies in it, this allows us to assume that n has no time component.

Subject to the condition of σ'' being flat[11], then, we can work in a reference frame where it is to be a surface of constant time, so that

$$\Phi\circ\Psi = \int d^Dx\,\Phi(\mathbf{x},t)^* i\overleftrightarrow{\partial_t}\Psi(\mathbf{x},t)\bigg|_{t=t''}. \tag{II.A.12}$$

The definition (II.6.14) of ρ_- means that

$$\Upsilon_{\lambda_2}\circ\rho_-\circ\Upsilon_{\lambda_1} = \int\frac{d^Dp}{(2\pi)^D 2\omega_{\mathbf{p}}}\left(e^{i\mathbf{p}\cdot\mathbf{x}}e^{i\omega_{\mathbf{p}}t}\circ\Upsilon_{\lambda_1}\right)\left(e^{i\mathbf{p}\cdot\mathbf{x}}e^{i\omega_{\mathbf{p}}t}\circ\Upsilon_{\lambda_2}\right)^*. \tag{II.A.13}$$

Now,

$$e^{i\mathbf{p}\cdot\mathbf{x}}e^{i\omega_{\mathbf{p}}t}\circ\Upsilon_\lambda = \int\frac{d^Dk}{(2\pi)^{D/2}\sqrt{2\omega_{\mathbf{k}}}}\widetilde{\Upsilon}^+(\mathbf{k})(\omega_{\mathbf{k}}-\omega_{\mathbf{p}})e^{-i(\omega_{\mathbf{k}}+\omega_{\mathbf{p}})t''}\int d^Dx\Theta(\lambda x_n)e^{i(\mathbf{k}-\mathbf{p})\cdot\mathbf{x}}; \tag{II.A.14}$$

the integral over $\mathbf{x}_\perp$ gives $(2\pi)^{D-1}\delta^{D-1}(\mathbf{k}_\perp-\mathbf{p}_\perp)$, and the integral over x_n gives

$$\begin{aligned}\int_{-\infty}^{\infty}dx_n\Theta(\lambda x_n)e^{i(k_n-p_n)x_n} &= \int_{-\lambda\infty}^{\lambda\infty}\lambda\,dx_n\Theta(x_n)e^{i\lambda(k_n-p_n)x_n} = \int_0^\infty dx_n e^{i\lambda(k_n-p_n)x_n}\\ &= \frac{i}{\lambda(k_n-p_n)+i\varepsilon} = \frac{i\lambda}{k_n-p_n}+\pi\delta(k_n-p_n).\end{aligned} \tag{II.A.15}$$

Substituting into (II.A.14) gives

$$e^{i\mathbf{p}\cdot\mathbf{x}}e^{i\omega_{\mathbf{p}}t}\circ\Upsilon_\lambda = i\lambda\int\frac{dk_n}{\sqrt{2\omega_{\mathbf{k}}}}(2\pi)^{D/2-1}\widetilde{\Upsilon}^+(\mathbf{k})\left(\frac{\omega_{\mathbf{k}}-\omega_{\mathbf{p}}}{k_n-p_n}\right)e^{-i(\omega_{\mathbf{k}}+\omega_{\mathbf{p}})t''} \tag{II.A.16}$$

[11]Note that if $D=1$, $\sigma''\cap\partial S(n)$ is a point, so this holds trivially.

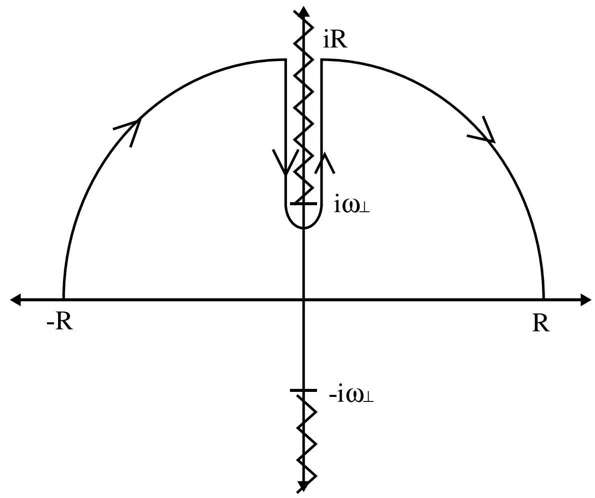


Figure II.9: The contour on which the integral in (II.A.17) is calculated to give (II.A.19). The radius R of the quarter-circle arcs is to be taken to infinity.

with $\mathbf{k}_\perp = \mathbf{p}_\perp$. We thus have

$$\begin{aligned}\Upsilon_{\lambda_2} \circ \rho_- \circ \Upsilon_{\lambda_1} = \lambda_1\lambda_2 \int \frac{dk_{1n}dk_{2n}d^{D-1}p_\perp}{2\sqrt{\omega_1\omega_2}(2\pi)^2} \widetilde{\Upsilon}^+(\mathbf{k}_2)^* \widetilde{\Upsilon}^+(\mathbf{k}_1) e^{-i(\omega_1-\omega_2)t''} \\ \times \int_{-\infty}^{\infty} \frac{dp_n}{2\omega_{\mathbf{p}}} \left(\frac{\omega_{\mathbf{p}}-\omega_1}{p_n - k_{1n}}\right)\left(\frac{\omega_{\mathbf{p}}-\omega_2}{p_n - k_{2n}}\right),\end{aligned} \tag{II.A.17}$$

where $\mathbf{k}_{1\perp} = \mathbf{k}_{2\perp} = \mathbf{p}_\perp$ so that $\omega_1 = \sqrt{k_{1n}^2 + \omega_\perp^2}$ and $\omega_2 = \sqrt{k_{2n}^2 + \omega_\perp^2}$ where $\omega_\perp = \sqrt{\mathbf{p}_\perp^2 + m^2}$. The integrand of the p_n integral,

$$f(p_n) = \frac{1}{2\omega_{\mathbf{p}}} \left(\frac{\omega_{\mathbf{p}}-\omega_1}{p_n - k_{1n}}\right)\left(\frac{\omega_{\mathbf{p}}-\omega_2}{p_n - k_{2n}}\right) \tag{II.A.18}$$

is analytic (since the singularities at $p_n = k_{1n}$ and $p_n = k_{2n}$ are removable) except for branch points when $\omega_{\mathbf{p}} = 0$, namely at $p_n = i\omega_\perp$ and $p_n = -i\omega_\perp$. We can thus deform the integration contour to the one shown in Fig. II.9. The contributions from the quarter-circle arcs cancel, and the contributions from the branch cut give

$$\int_{-\infty}^{\infty} f(p_n)dp_n = \int_{\omega_\perp}^{\infty} \frac{d\kappa}{\sqrt{\kappa^2 - \omega_\perp^2}} \frac{\omega_1\omega_2 + \omega_\perp^2 - \kappa^2}{(i\kappa - k_{1n})(i\kappa - k_{2n})}. \tag{II.A.19}$$

With the substitution $\kappa = \omega_\perp \sec\theta$, this becomes

$$\int_0^{\pi/2} \frac{\cos\theta(\omega_1\omega_2 + \omega_\perp^2 - \omega_\perp^2 \sec^2\theta)}{(i\omega_\perp - k_{1n}\cos\theta)(i\omega_\perp - k_{2n}\cos\theta)} d\theta, \tag{II.A.20}$$

which can be evaluated to give

$$\int_{-\infty}^{\infty} f(p_n)dp_n = \int_0^{\pi/2} \sec\theta d\theta + \frac{\omega_1+\omega_2}{k_{1n}-k_{2n}} \ln\left(\frac{\omega_1-k_{1n}}{\omega_2-k_{2n}}\right). \tag{II.A.21}$$

The first term is a constant, and is thus even in k_{1n}. The rest of (II.A.17) is odd in k_{1n} because of (II.A.11) so the constant term gives no contribution to $\Upsilon_{\lambda_2} \circ \rho_- \circ \Upsilon_{\lambda_1}$, and

$$\begin{aligned}
&\Upsilon_{\lambda_2} \circ \rho_- \circ \Upsilon_{\lambda_1} \\
&\quad = \lambda_1\lambda_2 \int \frac{dk_{1n}dk_{2n}d^{D-1}k_\perp}{(2\pi)^2} \frac{\omega_1+\omega_2}{2\sqrt{\omega_1\omega_2}} \widetilde{\Upsilon}^+(\mathbf{k}_2)^* \widetilde{\Upsilon}^+(\mathbf{k}_1) \frac{e^{-i(\omega_1-\omega_2)t''}}{k_{1n}-k_{2n}} \ln\left(\frac{\omega_1-k_{1n}}{\omega_2-k_{2n}}\right).
\end{aligned} \tag{II.A.22}$$

This gives us (II.7.1).

Chapter III

Non-Abelian Gauge Theories

III.1 Introduction

As discussed in Chapter I, General relativity (GR) possesses the symmetry of diffeomorphism invariance. In the $3 + 1$ formulation, this divides into time reparametrization invariance and the non-Abelian gauge group of spatial diffeomorphisms [33]. A technique for formulating quantum gravity, such as a generalized quantum mechanics defined by a sum over histories, will have to address the issues raised by these invariances, such as how (or whether) to enforce the constraints which the invariances imply. It is thus useful to examine proposed quantum formulations of GR by considering simpler theories exhibiting a subset of these invariances or similar ones. As described in Sec. I.4.6, the theory of a single relativistic particle exhibits reparametrization invariance similar to that of GR, while the spatial gauge transformations of GR are modelled by those of gauge theory. Previous work has applied the generalized quantum mechanics program to Abelian gauge theory [24]. In this chapter, I formulate a generalized quantum mechanics of a *non*-Abelian gauge theory, and examine the predictions for some sets of alternatives. That is to say, I construct exhaustive sets of mutually exclusive classes which partition the possible histories of the system, and for each such set of classes (or alternatives), I construct a decoherence functional. From this decoherence functional I determine whether probabilities can be defined for the set of alternatives, and if so what those probabilities are.

The role of this work with regard to the vast body of knowledge on Yang-Mills or non-Abelian gauge theories (see [34] for a review) is twofold: First, this is the first application of a "decoherence functional" or "consistent histories" method to their quantization.[1] As such, the focus is not primarily upon using such a theory for the practical consideration of the strong or weak interaction, but as a toy model which exhibits some features of GR. However, even as a quantization of a non-Abelian gauge theory itself, both the generalized

[1] Recent work [35] considers decoherence effects in the quantum cosmology of *massive* gauge fields, However, that work differs from the present enterprise in that the gauge fields are there seen as a model of matter coupled to gravity, while I consider massless gauge fields as a toy model for vacuum gravity itself. Even more significant is that while they study decoherence *effects*, it is in the context of a WKB quantization scheme, rather than a generalized quantum mechanics or consistent histories approach.

quantum mechanics formalism and this implementation thereof deal with different aspects of the theory than are usually considered. The alternatives for which generalized quantum mechanics predicts probabilities are not limited to projections onto eigenstates of operators at a single moment of time, but include alternatives defined by field averages over space*time* regions, which are inaccessible in a theory based on states and wave function reduction. This broader class of alternatives is especially of interest in connection with GR, where it is undesirable to single out a particular time variable for a conventional quantization. In addition, the considerations herein are predominantly nonperturbative, as contrasted with the usual perturbative scattering problems addressed in most practical treatments of Yang-Mills theory. On the other hand, the rich subject of topological aspects of non-Abelian gauge theories is not considered, and any potential global properties are in fact ignored by our assumptions about the behavior of fields at spatial infinity.

A second accomplishment of this work is that technical aspects of the path integrals involved in quantizing a non-Abelian gauge theory are more carefully considered than in the standard literature. Delicate issues involved in the time slicing of an explicit ("skeletonized") construction of the Lorentzian path integral (sections III.3 and III.4.3 are dealt with which are described only formally or implicitly in standard treatments such as [25]. Also, section III.6 exhibits a formal description of this quantization scheme which is manifestly Lorentz invariant.

The plan of this chapter is as follows: The review of non-Abelian gauge theory (NAGT) in Section III.2 establishes the perspective and notational conventions for the rest of the chapter, and describes the heuristic recipe for a generalized quantum mechanics of a NAGT.

Section III.3 describes the explicit implementation of that application, both as a formal path integral and in a spacetime lattice approximation[2] to the path integral. Technology is developed therein for handling the lattice expressions (in particular the time slicing) explicitly, which should be of use in other treatments of path integrals as well. In the latter half of the section, we verify explicitly that the implementation is gauge invariant. We also show there that our sum-over-histories expression agrees, in its description of the propagator, with the results of a reduced phase space canonical operator theory in which the constraints (Gauss's law) are enforced before quantization. In our generalized quantum mechanics formulation, the constraints are not enforced identically–as they are in a reduced phase space implementation–but are quantities whose values must be predicted by the theory. Thus the prediction of probabilities for the values of the constraints occupies most of our attention in the remainder of the chapter.

Section III.4 considers one subset of all possible alternatives which defines a "phase space" realization (as defined in section III.4.1) of the physical gauge fields. The predictions of such a theory are found to be consistent with the vanishing of the constraints for nearly all such sets of alternatives, and thus to agree with those of a reduced phase space canonical theory. In section III.5 we consider another subset of the allowed alternatives in which the gauge electric field is realized in terms of the potentials rather than their conjugate momenta. Since the momenta are then not restricted by the alternatives, we perform the integrals over

[2]This is not in the sense of lattice gauge theory with its Euclidean lattice, Wilson loops, *etc.,* but simply a means to provide what Hatfield [36] calls a constructive definition of the path integral.

them and reduce our theory to a "configuration space" one. Now, defining the constraints by their configuration space realizations, we find the two most significant results of the chapter. In section III.5.3 we see that for some quantities which vanish when the constraints are satisfied we recover the result of [24] for electromagnetism, namely, either the quantities vanish with probability one or quantum mechanical interference prevents us from assigning probabilities to possible outcomes. We also verify that there are coarse grainings which fall into the first category. However, the result does not necessarily hold for all quantities which vanish in the presence of the constraints, and in section III.5.4 we exhibit such a quantity in the Abelian theory of electromagnetism for which we predict a nonzero probability of an alternative inconsistent with the constraints. Since this set of alternatives involves averages of fields over time, it is not accessible in less general quantum Yang-Mills theories. From a spacetime point of view, it is sensible that the constraints do not have a special status in this theory, being just one component of the equations of motion.

Finally, section III.6 verifies that the configuration space theory is Lorentz-invariant by casting the formal path integral in a form where that invariance is manifest, even in the attachment of the initial and final states.

III.2 Formulation

III.2.1 Non-Abelian gauge theories

Fields

In this section we set out the conventions used herein to describe a non-Abelian gauge theory (NAGT) in flat spacetime with the metric diag$(-1, 1, 1, 1)$.

The gauge group is described by Hermitian generators $\{T_a\}$ with real, totally anti-symmetric structure constants $\{f^c_{ab}\}$: $[T_a, T_b] = if^c_{ab}T_c$.

A gauge transformation is described by a matrix $U = e^{ig\Lambda_a(x)T_a}$. The connection is a four-vector $A^a(x)$ with components $\{A^a_\mu(x)\}$, which transforms under infinitesimal gauge transformations according to

$$\delta A^a_\mu = -\nabla_\mu \,\delta\Lambda_a - gf^c_{ab}A^c_\mu \delta\Lambda^b. \tag{III.2.1}$$

If we define a covariant derivative

$$D_\mu = \nabla_\mu + igA^a_\mu T_a, \tag{III.2.2}$$

it transforms according to

$$D_\mu \to UD_\mu U^{-1}. \tag{III.2.3}$$

This means that if ψ is an isovector, i.e., a vector in the same space as the matrices $\{T_a\}$ which transforms under gauge transformations according to $\psi \to U\psi$, the covariant gradient of ψ will transform the same way:

$$D_\mu\psi(x) \to UD_\mu\psi(x). \tag{III.2.4}$$

The field strength tensor is

$$G^a_{\mu\nu} T_a = \frac{[D_\mu, D_\nu]}{ig}. \tag{III.2.5}$$

In a particular Lorentz frame, we divide the connection A^a into scalar and vector potentials φ^a and $\mathbf{A}^a$:

$$\varphi_a = A^0_a \tag{III.2.6a}$$
$$\mathbf{A}_a = A^i_a \mathbf{e}_i \tag{III.2.6b}$$

and the field strength tensor into gauge electric and magnetic fields $\mathbf{E}^a$ and $\mathbf{B}^a$:

$$E^i_a = G^{0i}_a \tag{III.2.7a}$$
$$B^i_a = \frac{1}{2}\epsilon^{ijk} G^{jk}_a \tag{III.2.7b}$$

where ϵ^{ijk} is the Levi-Civita symbol. The gauge electric and magnetic fields can then be expressed in terms of the scalar and vector potentials as

$$\mathbf{E}_a = -\dot{\mathbf{A}}_a - \boldsymbol{\nabla}\varphi_a - g f^c_{ab} \mathbf{A}_c \varphi_b \tag{III.2.8a}$$
$$\mathbf{B}_a = \boldsymbol{\nabla} \times \mathbf{A}_a + \frac{1}{2} g f^c_{ab} \mathbf{A}_c \times \mathbf{A}_b. \tag{III.2.8b}$$

The gauge electric and magnetic fields can be shown to transform under gauge transformations as follows:

$$\mathbf{E}_a T_a \to U \mathbf{E}_a T_a U^{-1} \tag{III.2.9a}$$
$$\mathbf{B}_a T_a \to U \mathbf{B}_a T_a U^{-1}, \tag{III.2.9b}$$

which becomes, for an infinitesimal transformation,

$$\delta\mathbf{E}_a = -g f^c_{ab} \delta\Lambda_b \mathbf{E}_c \tag{III.2.10a}$$
$$\delta\mathbf{B}_a = -g f^c_{ab} \delta\Lambda_b \mathbf{B}_c. \tag{III.2.10b}$$

This is the transformation property of an isovector in the adjoint representation [37], in which the generators are represented by $(T^c)_{ab} = -i f^c_{ab}$, so we will often drop the index from $\mathbf{E}$ or $\mathbf{B}$ and consider it to be an isovector in the adjoint representation. The connection A^a has an inhomogeneous piece in its transformation law (III.2.1), so it is not a true isovector, but we will represent it as one notationally. Thus the gauge electric field can be written as[3]

$$\mathbf{E} = -\dot{\mathbf{A}} - \mathbf{D}\varphi, \tag{III.2.11}$$

where we have realized $\mathbf{D}$ in the adjoint representation as

$$\mathbf{D}_{ab} = \delta_{ab} \boldsymbol{\nabla} + g f^c_{ab} \mathbf{A}_c. \tag{III.2.12}$$

[3]Neither $\dot{\mathbf{A}} = \partial_t \mathbf{A}$ nor $\mathbf{D}\varphi$ alone is an isovector, but their sum $-\mathbf{E}$ is.

Classical equations of motion

The action for a NAGT in the absence of matter is

$$S = \int d^4x \mathcal{L} = -\int d^4x \frac{1}{4} G_a^{\mu\nu} G_{\mu\nu}^a = \int d^4x \frac{1}{2}(\mathbf{E}^2 - \mathbf{B}^2). \qquad \text{(III.2.13)}$$

The conjugate momenta in a particular reference frame are found by differentiating the Lagrangian density $\mathcal{L}$ with respect to $\dot{A}^\mu = \partial_t A^\mu$:

$$\pi_0 = \frac{\mathcal{D}\mathcal{L}}{\mathcal{D}\dot{\varphi}} = 0 \qquad \text{(III.2.14a)}$$

$$\boldsymbol{\pi} = \frac{\mathcal{D}\mathcal{L}}{\mathcal{D}\dot{\mathbf{A}}} = \dot{\mathbf{A}} + \mathbf{D}\varphi = -\mathbf{E}. \qquad \text{(III.2.14b)}$$

The Hamiltonian density is given by

$$\mathcal{H}[A, \boldsymbol{\pi}] = \dot{\mathbf{A}} \cdot \boldsymbol{\pi} - \mathcal{L} = \frac{1}{2}\boldsymbol{\pi}^2 + \frac{1}{2}\mathbf{B}^2 - \boldsymbol{\pi} \cdot \mathbf{D}\varphi = \mathcal{H}[\mathbf{A}, \boldsymbol{\pi}] - \boldsymbol{\pi} \cdot \mathbf{D}\varphi \qquad \text{(III.2.15)}$$

and Hamilton's equations of motion are

$$\dot{\mathbf{A}} = \boldsymbol{\pi} - \mathbf{D}\varphi \qquad \text{(III.2.16a)}$$

$$D_t \boldsymbol{\pi} = \mathbf{D} \times \mathbf{B} \qquad \text{(III.2.16b)}$$

$$\mathbf{D} \cdot \boldsymbol{\pi} = 0. \qquad \text{(III.2.16c)}$$

Equation (III.2.16c) involves no time derivatives, so it is the constraint of this NAGT, which we call $K = \mathbf{D} \cdot \boldsymbol{\pi}$.

III.2.2 Generalized quantum mechanics applied to a NAGT

To formulate a NAGT in generalized quantum mechanics, we follow a procedure similar to the one described in [24] for electromagnetism (E&M).

To ensure that the theory is sufficiently general to allow coarse graining by values of the constraint, we work in the sum-over-histories formulation described in Sec. I.3.3, in which the initial state is described by a set of wave functionals $\{\Psi_j[A']\}$ with corresponding non-negative weights $\{p'_j\}$ and the final state by a set of wave functionals $\{\Phi_i[A'']\}$ and weights $\{p''_i\}$, and the decoherence functional is given by

$$D(\alpha, \alpha') = \frac{\sum_{i,j} p''_i \langle \Phi_i | C_\alpha | \Psi_j \rangle \langle \Phi_i | C_{\alpha'} | \Psi_j \rangle^* p'_j}{\sum_{i,j} p''_i \left| \langle \Phi_i | C_u | \Psi_j \rangle \right|^2 p'_j}. \qquad \text{(I.3.12)}$$

The quantity $\langle \Phi_i | C_\alpha | \Psi_j \rangle$ corresponding to a matrix element of the class operator for the class c_α, is constructed by a sum over the histories in the class c_α, weighted by the initial and final wave functionals Ψ_j and Φ_i. Schematically (see (I.3.13)):

$$\langle \Phi_i | C_\alpha | \Psi_j \rangle = \sum_{\text{history} \in \alpha} \Phi_i^*[\text{endpt}''] e^{iS[\text{history}]} \Psi_j[\text{endpt}']. \qquad \text{(III.2.17)}$$

The wave functionals[4] are taken to be functionals of scalar and vector potential configurations on an initial or final surface of constant time, as appropriate. They are assumed to obey the operator form of the constraints $\pi_0 = 0$ and $\mathbf{D} \cdot \boldsymbol{\pi} = 0$:

$$\frac{\mathcal{D}}{\mathcal{D}\varphi'} \Psi[A'; t') = 0 \tag{III.2.18a}$$

$$\mathbf{D} \cdot \frac{\mathcal{D}}{\mathcal{D}\mathbf{A}'} \Psi[A'; t') = 0 \tag{III.2.18b}$$

and likewise for $\Phi[A''; t'')$. (Since the wave functionals are independent of the scalar potential φ, we will henceforth write the first argument as the three-vector $\mathbf{A}'$ rather than the four-vector A'.) Continuing the analogy between GR and a NAGT begun in Sec. I.4.6, (III.2.18b) is analogous to the momentum constraint in GR, while (III.2.18a) corresponds to the lack of dependence of the wave functional for quantum GR on the shift vector $\{N^i\}$.

Having described schematically the construction of the decoherence functional, we now specify the other two elements which describe the generalized quantum mechanics. The fine grained histories summed over are complete field configurations $A(x)$ [and also $\boldsymbol{\pi}(x)$ if we are considering a phase space formulation] in the region between the initial and final time slices. The allowable coarse grainings are limited to gauge invariant partitions of the fields.

III.3 Class operators in the path integral formulation

III.3.1 Overview

This section describes in detail how to implement the sum over histories heuristically described in (III.2.17). In section III.3.2 we express this as a formal path integral. Section III.3.3 contains an explicit realization of this integral on a discrete spacetime lattice, where the lattice spacing is to be taken to be infinitesimally small.[5] The following two sections demonstrate that the particular details chosen in section III.3.3 were suitable by showing that the path integral has desired properties. In section III.3.4, the sum-over-histories expression for the class operator C_u corresponding to the class c_u of all paths is shown to equal, up to a constant multiplicative factor, the propagator $e^{i\widehat{H}_{\mathrm{red}}T}$ in a reduced phase space canonical theory. In section III.3.5 the path integral is shown to be unchanged under the discrete equivalent of a gauge transformation, in the limit that the lattice spacing goes to zero.

[4] Recall that the conventions defined in Sec. I.A indicate that $\Psi[A', t')$ is a functional of field configurations $\{A'(\mathbf{x})\}$ and a function of time t'.

[5] It should be stated once again that we are *not* doing lattice gauge theory in anything like the usual sense. The action is expressed directly in terms of fields defined at each lattice point and not in terms of the "links" defining a "plaquette". In addition, our spacetime lattice is Lorentzian rather than Euclidean.

III.3.2 Formal expression

We express the sum over histories (III.2.17) via a path integral which is formally written as[6]

$$\langle\Phi|C_\alpha|\Psi\rangle = \int_\alpha \mathcal{D}^4A\mathcal{D}^3\pi\Phi^*[\mathbf{A}'';t'']\,\delta[G]\Delta_G[A,\boldsymbol{\pi}]e^{iS_{\text{can}}[A,\boldsymbol{\pi}]}\Psi[\mathbf{A}';t'], \tag{III.3.1}$$

where the gauge condition is $G=0$ and Δ_G is the corresponding Fadeev-Popov gauge-fixing determinant. It was originally defined in [25] in terms of the Poisson bracket:

$$\Delta_G = |\det\{G,K\}|\,. \tag{III.3.2a}$$

Other useful (and equivalent) definitions are (see e.g., [36])

$$\Delta_G = \left|\det\left[\frac{\mathcal{D}G^\Lambda}{\mathcal{D}\Lambda}\right]\right| \tag{III.3.2b}$$

and

$$\frac{1}{\Delta_G} = \int \mathcal{D}\Lambda\,\delta[G^\Lambda - G], \tag{III.3.2c}$$

where Λ is the parameter defining a gauge transformation which takes G into G^Λ.

Finally, the canonical action is

$$S_{\text{can}} = \int d^4x\left(\dot{\mathbf{A}}\cdot\boldsymbol{\pi} - \mathcal{H}[A,\boldsymbol{\pi}]\right) = \int d^4x\left(\dot{\mathbf{A}}\cdot\boldsymbol{\pi} - \frac{1}{2}\boldsymbol{\pi}^2 - \frac{1}{2}\mathbf{B}^2 + \boldsymbol{\pi}\cdot\mathbf{D}\varphi\right). \tag{III.3.3}$$

If we assume[7] that φ vanishes at spatial infinity, we can integrate by parts[8] to obtain

$$S_{\text{can}} = \int d^4x\left(\dot{\mathbf{A}}\cdot\boldsymbol{\pi} - \frac{1}{2}\boldsymbol{\pi}^2 - \frac{1}{2}\mathbf{B}^2 - \varphi\mathbf{D}\cdot\boldsymbol{\pi}\right). \tag{III.3.4}$$

The expression (III.3.1) involves the full set of phase space variables, but we will also use it as the starting point for the configuration space formulation. If our coarse graining makes no reference to the conjugate momentum $\boldsymbol{\pi}$, we can work in a gauge which does not restrict $\boldsymbol{\pi}$ and integrate it out to obtain

$$\langle\Phi|C_\alpha|\Psi\rangle = \int_\alpha \mathcal{D}^4A\,\Phi^*[\mathbf{A}'';t'']\delta[G]\Delta_G[A]e^{iS[A]}\Psi[\mathbf{A}';t'], \tag{III.3.5}$$

where S is the (configuration space) action (III.2.13) and for the purposes of this formal expression, a constant factor has been absorbed into $\mathcal{D}^4A$.

---

[6] For those wondering what has become of the generic inner product $\circ$ from (I.3.13a), the inner product necessary to produce states with finite norm is one which includes a gauge-fixing mechanism, as in (III.3.9). In the expression (III.3.1), the gauge-fixing factors $\delta[G']\Delta'_G$ and $\delta[G'']\Delta''_G$ have been absorbed into the factor $\delta[G]\Delta_G$ pertaining to the entire path.

[7] Throughout this chapter we will neglect any global issues such as the Gribov ambiguity [38] and assume that fields can be taken to vanish at spatial infinity.

[8] It is worth pointing out once explicitly that the covariant gradient $\mathbf{D}$ behaves like the ordinary gradient under integration by parts. Examining $\alpha\mathbf{D}\beta = \alpha_a\boldsymbol{\nabla}\beta_a + f^c_{ab}\alpha_a\beta_b\mathbf{A}_c$, we see that the first term integrates by parts as usual, and the second term also picks up a minus sign under the interchange of α and β due to the antisymmetry of f^c_{ab}. Thus $\alpha\mathbf{D}\beta = \boldsymbol{\nabla}(\alpha\beta) - \beta\mathbf{D}\alpha$.

III.3.3 Lattice realization

To give a concrete meaning to the formal path integral in (III.3.1), we imagine it to be defined on an arbitrarily small lattice; the spatial volume is divided into lattice elements of volume $\delta^3 x$ and the time interval from t' to t'' is divided into slices of separation $\delta t = \frac{t''-t}{J+1} \equiv \frac{T}{J+1}$. The lattice expression for the class operator is then

$$\begin{aligned}\langle\Phi|C_\alpha|\Psi\rangle = \int \mathcal{D}^4 A^{J+1} \Phi^*\left[\mathbf{A}^{J+1};t''\right) \times \\ \left(\prod_{M=J}^{0} \int \mathcal{D}^4 A^M \mathcal{D}^3 \pi^M \exp\left\{i\delta t \int d^3x \left(\dot{\mathbf{A}}^M \cdot \boldsymbol{\pi}^M - \mathcal{H}\left[\overline{\mathbf{A}}^M, \boldsymbol{\pi}^M\right] - \overline{\varphi}^M \overline{\mathbf{D}}^M \cdot \boldsymbol{\pi}^M\right)\right\}\right) \\ \times \left(\prod_{M=0}^{J+2} \delta\left[G^M\right] \Delta_{G^M}\right) \Psi\left[\mathbf{A}^0;t'\right) e_\alpha\left[A,\boldsymbol{\pi}\right],\end{aligned} \tag{III.3.6}$$

where the "barred" quantities indicate temporal averages:

$$\overline{A}^M = \frac{A^{M+1}+A^M}{2}, \quad 0 \le M \le J \tag{III.3.6a}$$

$$\overline{\mathbf{D}}^M_{ab} = \delta_{ab}\boldsymbol{\nabla} + g f^c_{ab} \overline{\mathbf{A}}^M_c \tag{III.3.6b}$$

and the lattice expression for the "velocity" is

$$\dot{\mathbf{A}}^M = \frac{\mathbf{A}^{M+1}-\mathbf{A}^M}{\delta t}, \quad 0 \le M \le J. \tag{III.3.6c}$$

The expression (III.3.6) reflects the fact that since the "velocity" $\dot{\mathbf{A}}^M$ is naturally associated with a point *halfway* between the coordinate lattice slices labeled M and $M+1$ by (III.3.6c), it is sensible to associate the conjugate momenta $\boldsymbol{\pi}^M$ with those points as well, in light of the term $\dot{\mathbf{A}}^M \cdot \boldsymbol{\pi}^M$. This means that to maintain manifest time reversal symmetry we should not associate $\boldsymbol{\pi}^M$ with A^M or A^{M+1}, but instead with $\overline{A}^M$. The gauge-fixing expressions G^M and Δ_{G^M} are also assumed to be expressed in terms of the averaged fields $\left\{\overline{A}^M\right\}$ whenever their complexity prevents an unambiguous definition in terms of the $\left\{A^M\right\}$ alone.[9]

The factor of e_α is a functional of the paths which is unity for any path in the class c_α and vanishes for any path not in c_α.

The remaining functional integrals $\left\{\mathcal{D}^4 A^M\right\}$ and $\left\{\mathcal{D}^3 \pi^M\right\}$ are over functions of the spatial coördinate $\mathbf{x}$. We leave consideration of the spatial dependence somewhat formal, because all of the complications involved in describing the gradients on the lattice basically appear, and more seriously, in the temporal direction. We will thus be speaking as though

[9] The reason why there are $J+3$ gauge conditions $\left\{G^M\right\}$ is most easily seen in the temporal gauge $\varphi \equiv 0$. On a lattice this corresponds to the $J+2$ conditions $\left\{\varphi^M = 0, 0 \le M \le J+1\right\}$. However, there is residual gauge freedom in the temporal gauge, since a gauge transformation by a parameter $\Lambda(\mathbf{x})$ [see (III.2.1)] which is independent of the time t will preserve the temporal gauge condition $\varphi \equiv 0$. To completely fix the gauge, then, we would need to specify one other quantity over all space at a particular time. This is the last of the $J+3$ gauge conditions.

the spacetime function $A(x)$ is broken up into a series of functions of a continuous spatial variable $\mathbf{x}$: $\{A^M(\mathbf{x})\}$. However, we can ultimately consider the following lattice resolutions for the functional integrals and delta functions:

$$\mathcal{D}A^{M\mu} = \prod_{a,\mathbf{x}} N_A dA_a^{M\mu}(\mathbf{x}) \tag{III.3.7a}$$

$$\mathcal{D}\pi^{Mi} = \prod_{a,\mathbf{x}} N_\pi d\pi_a^{Mi}(\mathbf{x}) \tag{III.3.7b}$$

$$\delta\left[A^{M\mu}\right] = \prod_{a,\mathbf{x}} \delta\big(A_a^{M\mu}(\mathbf{x})\big) \Big/ N_A \tag{III.3.7c}$$

$$\delta\left[\pi^{Mi}\right] = \prod_{a,\mathbf{x}} \delta\big(\pi_a^{Mi}(\mathbf{x})\big) \Big/ N_\pi, \tag{III.3.7d}$$

where N_A and N_π are arbitrary normalization constants which obey

$$N_A N_\pi = \frac{\delta^3 x}{2\pi}. \tag{III.3.8}$$

{This definition is chosen to give the desirable properties (III.3.27) for potential and momentum eigenstates in the corresponding operator theory. See also equation (6.228) of [39].}

Since (III.3.8) only defines a relation between N_A and N_π, there is still an arbitrary factor in the definitions of the measures (III.3.7). It is reassuring to verify that the class operator matrix elements $\langle\Phi|C_\alpha|\Psi\rangle$ defined by (III.3.6) are independent of that arbitrary factor so that for instance if we double N_A (and thus halve N_π), they are unchanged. To do this correctly, we must also keep in mind that the wave functionals $\Psi[\mathbf{A}']$ and $\Phi[\mathbf{A}'']$ also depend on the value of N_A as follows. Let the wave functionals be normalized according to[10]

$$\int \mathcal{D}^3 A' \delta[G[\mathbf{A}']]\Delta_G[\mathbf{A}'] \left|\Psi\left[\mathbf{A}'\right]\right|^2 = 1. \tag{III.3.9}$$

Since there is a factor of $\mathcal{N}_A^3$ (where $\mathcal{N}_A = \prod_{a,\mathbf{x}} N_A$ and $\mathcal{N}_\pi = \prod_{a,\mathbf{x}} N_\pi$) associated with the measure $\mathcal{D}^3 A'$ and a factor of $\mathcal{N}_A^{-1}$ associated with the gauge fixing delta function $\delta[G']$, there must be a factor of $\mathcal{N}_A^{-2}$ associated with the square of the wave functional so that the factors all cancel out. That is, if we double $\mathcal{N}_A$, we must halve $\Psi[\mathbf{A}']$ to maintain the normalization (III.3.9). Considering the expression (III.3.6), if we multiply together all the normalization factors ($\mathcal{N}_A$ for each $\mathcal{D}A^{M\mu}$, $\mathcal{N}_\pi$ for each $\mathcal{D}\pi^{Mi}$, $\mathcal{N}_A^{-1}$ for each $\delta[G_M]$, and $\mathcal{N}_A^{-1}$ for each wave functional) we have

$$\mathcal{N}_A^4 \mathcal{N}_A^{-1} (\mathcal{N}_A^4 \mathcal{N}_\pi^3)^{J+1} (\mathcal{N}_A^{-1})^{J+3} \mathcal{N}_A^{-1} = \mathcal{N}_A^{3J+3} \mathcal{N}_\pi^{3J+3} = \prod_{a,\mathbf{x}} \left(\frac{\delta^3 x}{2\pi}\right)^{3J+3} \tag{III.3.10}$$

and the arbitrariness of the normalization N_A does indeed cancel out of (III.3.6).

[10] The presence of the $\delta[G']\Delta'_G$ is necessary to produce a finite inner product $\Psi \circ \Psi$; see footnote 6 on page 65.

III.3.4 Relation to reduced phase space theory

To illustrate why the details of (III.3.6) were chosen, we show in this section that it gives the same expression, up to a normalization constant, for the propagator C_u as would be obtained from a canonical theory working with only the "physical degrees of freedom". [Comparisons of (III.3.6) to the reduced phase space results for particular coarse grainings will be considered in later sections.] This derivation is essentially the one given in [25], taken in reverse order and with more attention paid to the details of the lattice.

First, we observe that there are $J+2$ coördinate variables $\left\{A^M(\mathbf{x})\right\}$ but only $J+1$ averaged variables $\left\{\overline{A}^M(\mathbf{x})\right\}$. We can define $\overline{A}^{J+1}$ to be a linear combination of the $\left\{A^M\right\}$ independent of the $J+1$ other $\left\{\overline{A}^M\right\}$, normalized so that the change of variables from $\left\{A^M\right\}$ to $\left\{\overline{A}^M\right\}$ has unit determinant and

$$\prod_{M=0}^{J+1} dA_a^{M\mu}(\mathbf{x}) = \prod_{M=0}^{J+1} d\overline{A}_a^{M\mu}(\mathbf{x}). \tag{III.3.11}$$

{One such choice is $\overline{A}^{J+1} = 2^J\left[A^{J+1} + (-1)^{J+1}A^0\right]$.} Since neither the velocity term $\dot{\mathbf{A}}\cdot\boldsymbol{\pi}$ nor the initial and final wave functionals depends on the scalar potential φ, it only enters (III.3.6) through the "genuine" averages $\left\{\overline{\varphi}^M, 0 \leq M \leq J\right\}$, and thus the class operator is independent of $\overline{\varphi}^{J+1}$. It is then natural to choose $\overline{\varphi}^{J+1}(\mathbf{x}) = 0$ as one of our $J+3$ time slices worth of gauge conditions. The corresponding Fadeev-Popov determinant is

$$\Delta_{\overline{\varphi}^{J+1}} = \det[D_t]|_{\varphi=0} = \det[\partial_t], \tag{III.3.12}$$

which is a constant. Overall constant (i.e., the same for all α) factors in the class operator $\langle\Phi|C_\alpha|\Psi\rangle$ will cancel out in the expression (I.3.12) for the decoherence functional and will not affect the physics.

For this demonstration, it is simplest to choose as the bulk of the gauge conditions the axial gauge, in which the component of $\mathbf{A}$ along some fixed unit vector $\mathbf{e}_n$ vanishes:

$$\mathbf{e}_n\cdot\mathbf{A}(x) \equiv A_n(x) = 0. \tag{III.3.13}$$

The Fadeev-Popov determinant of this gauge condition is

$$\Delta_{A_n} = \det[-D_n]|_{A_n=0} = \det[-\partial_n], \tag{III.3.14}$$

another constant. The remaining components of the vector potential are

$$\mathbf{A}_\perp = \mathbf{A} - A_n\mathbf{e}_n. \tag{III.3.15}$$

Thus the class operator becomes

$$\langle\Phi|C_\alpha|\Psi\rangle = \int \mathcal{D}^2 A_\perp^{J+1} \Phi^*\left[\mathbf{A}_\perp^{J+1};t''\right) \det[\partial_t]\det[-\partial_n]\left(\prod_{M=J}^{0}\int \mathcal{D}^2 A_\perp^M \mathcal{D}\overline{\varphi}^M \mathcal{D}^3\pi^M\right.$$
$$\left.\times\exp\left\{i\delta t\int d^3x\left(\dot{\mathbf{A}}_\perp^M\cdot\boldsymbol{\pi}_\perp^M - \mathcal{H}\left[\overline{\mathbf{A}}_\perp^M,\boldsymbol{\pi}^M\right] - \overline{\varphi}^M\overline{\mathbf{D}}^M\cdot\boldsymbol{\pi}^M\right)\right\}\det[-\partial_n]\right)$$
$$\times\,\Psi\left[\mathbf{A}_\perp^0;t'\right)e_\alpha[A,\boldsymbol{\pi}]. \quad \text{(III.3.16)}$$

To specialize to the propagator C_u, which is defined by a sum over *all* paths, we set $e_\alpha = 1$. We can perform each of the $\overline{\varphi}$ integrals to obtain

$$\int \mathcal{D}\overline{\varphi}^M \exp\left(-i\delta t\int d^3x\,\overline{\varphi}^M\overline{\mathbf{D}}^M\cdot\boldsymbol{\pi}^M\right)$$
$$= \prod_{a,\mathbf{x}}\int N_A d\overline{\varphi}_a^M(\mathbf{x})\exp\left[-i\delta t\,\delta^3x\,\overline{\varphi}_a^M(\mathbf{x})\left(\overline{\mathbf{D}}^M\cdot\boldsymbol{\pi}^M\right)_a(\mathbf{x})\right]$$
$$= \prod_{a,\mathbf{x}} N_A\frac{2\pi}{\delta^3x\delta t}\delta\left(\left(\overline{\mathbf{D}}^M\cdot\boldsymbol{\pi}^M\right)_a(\mathbf{x})\right) = \delta\left[\overline{\mathbf{D}}^M\cdot\boldsymbol{\pi}^M\right]\delta T^{-1}, \quad \text{(III.3.17)}$$

where we have defined the infinite constant

$$\delta T = \prod_{a,\mathbf{x}}\delta t. \quad \text{(III.3.18)}$$

In a reduced phase space theory, the gauge component (here A_n) of the coördinate is taken to vanish, and the corresponding component of the conjugate momentum is restricted to the value which causes the constraint to be satisfied. [Since the Lagrange multipliers (here the scalar potential φ) multiply identically enforced constraints, they do not appear in the Lagrangian.] In the axial gauge, this means that π_n has the value which ensures $K = \mathbf{D}\cdot\boldsymbol{\pi} = 0$, or

$$D_n\pi_n = \partial_n\pi_n = -\mathbf{D}_\perp\cdot\boldsymbol{\pi}_\perp. \quad \text{(III.3.19)}$$

We define a functional which accomplishes that:

$$\pi_n\left[\mathbf{A}'_\perp,\boldsymbol{\pi}'_\perp;\mathbf{x}\right) = \int^{x_n} dx'_n(-\mathbf{D}'_\perp\cdot\boldsymbol{\pi}'_\perp)(\mathbf{x}_\perp+\mathbf{e}_n x'_n) = -\left(\partial_n^{-1}\mathbf{D}'_\perp\cdot\boldsymbol{\pi}'_\perp\right)(\mathbf{x}). \quad \text{(III.3.20)}$$

Returning to the sum-over-histories expression, we observe that the delta function from (III.3.17) enforcing the constraint can be rewritten

$$\delta\left[\overline{\mathbf{D}}^M\cdot\boldsymbol{\pi}^M\right] = \delta\Big[\pi_n^M - \pi_n\left[\overline{\mathbf{A}}_\perp^M,\boldsymbol{\pi}_\perp^M\right]\Big]\left|\det\left[\frac{\mathcal{D}\pi_n^M}{\mathcal{D}K^M}\right]\right|. \quad \text{(III.3.21)}$$

Recalling (III.3.2a) we identify the determinant in the expression above as the reciprocal of the Fadeev-Popov determinant, since

$$\Delta_{A_n} = \left|\det[\{A_n,K\}]\right| = \det\left[\frac{\mathcal{D}K}{\mathcal{D}\pi_n}\right] = \det[D_n]. \quad \text{(III.3.22)}$$

Thus we can combine the Fadeev-Popov determinant with the φ integral producing the delta function (III.3.21) to obtain

$$\delta\Big[\overline{A}_n^M\Big]\,\Delta_{\overline{A}_n^M}\int\mathcal{D}\overline{\varphi}^M\exp\left(-i\delta t\int d^3x\,\overline{\varphi}^M\overline{\mathbf{D}}^M\!\cdot\boldsymbol{\pi}^M\right)=\delta\Big[\overline{A}_n^M\Big]\,\delta\Big[\pi_n^M-\pi_n\left[\overline{\mathbf{A}}_\perp^M,\boldsymbol{\pi}_\perp^M\right]\Big]\delta T^{-1}, \tag{III.3.23}$$

which means we can rewrite the propagator as

$$\begin{aligned}\langle\Phi|C_u|\Psi\rangle=\int\mathcal{D}^2A_\perp^{J+1}\Phi^*\left[\mathbf{A}_\perp^{J+1};t''\right)\det[\partial_t]\det[-\partial_n]\delta T^{-1}\Bigg(\prod_{M=J}^{0}\int\mathcal{D}^2A_\perp^M\mathcal{D}^2\pi_\perp^M\\ \times\exp\left\{i\delta t\int d^3x\left(\dot{\mathbf{A}}_\perp^M\cdot\boldsymbol{\pi}_\perp^M-\mathcal{H}_{\rm red}\left[\overline{\mathbf{A}}_\perp^M,\boldsymbol{\pi}_\perp^M\right]\right)\right\}\Bigg)\Psi\left[\mathbf{A}_\perp^0;t'\right),\end{aligned} \tag{III.3.24}$$

where $\mathcal{H}_{\rm red}$ is the reduced Hamiltonian density

$$\mathcal{H}_{\rm red}[\mathbf{A}'_\perp,\boldsymbol{\pi}'_\perp]=\mathcal{H}\Big[\mathbf{A}'_\perp,\boldsymbol{\pi}'_\perp+\mathbf{e}_n\pi_n[\mathbf{A}'_\perp,\boldsymbol{\pi}'_\perp]\Big]. \tag{III.3.25}$$

To convert (III.3.24) into a canonical operator form of the propagator, we consider eigenstates of the “physical” coördinate and momentum operators:

$$\widehat{\mathbf{A}}_\perp|\mathbf{A}'_\perp\rangle=\mathbf{A}'_\perp|\mathbf{A}'_\perp\rangle,\qquad\widehat{\boldsymbol{\pi}}_\perp|\boldsymbol{\pi}'_\perp\rangle=\boldsymbol{\pi}'_\perp|\boldsymbol{\pi}'_\perp\rangle, \tag{III.3.26}$$

normalized so that

$$\langle\mathbf{A}''_\perp|\mathbf{A}'_\perp\rangle=\delta[\mathbf{A}''_\perp-\mathbf{A}'_\perp] \tag{III.3.27a}$$

$$\langle\mathbf{A}'_\perp|\boldsymbol{\pi}'_\perp\rangle=\exp\left(i\int d^3x\,\mathbf{A}'_\perp\cdot\boldsymbol{\pi}'_\perp\right), \tag{III.3.27b}$$

which implies, via (III.3.7)-III.3.8), that

$$\langle\boldsymbol{\pi}''_\perp|\boldsymbol{\pi}'_\perp\rangle=\delta[\boldsymbol{\pi}''_\perp-\boldsymbol{\pi}'_\perp]. \tag{III.3.27c}$$

A state vector corresponding to each wave functional is then defined by

$$|\Psi\rangle=\int\mathcal{D}^2A'_\perp|\mathbf{A}'_\perp\rangle\Psi[\mathbf{A}'_\perp]. \tag{III.3.28}$$

The operator form of the reduced Hamiltonian density has an ambiguity because of the operator ordering of the term $\frac{1}{2}\left(\pi_n[\mathbf{A}'_\perp,\boldsymbol{\pi}'_\perp]\right)^2$. A natural[11] choice is Weyl ordering

[11]Different choices of operator ordering will typically lead to different lattice realizations than the one in (III.3.6), although no systematic description of the correspondence is known to the author. Two simple examples of alternate operator ordering are one in which all the $\widehat{\mathbf{A}}$'s are placed to the right of all the $\widehat{\boldsymbol{\pi}}$'s and one in which the $\widehat{\boldsymbol{\pi}}$'s are to the right of the $\widehat{\mathbf{A}}$'s. The former will lead to a lattice expression in which $\boldsymbol{\pi}^M$ and $\dot{\mathbf{A}}^M$ are associated with A^M, while the latter will associate $\boldsymbol{\pi}^M$ and $\dot{\mathbf{A}}^M$ with A^{M+1}. Weyl ordering is preferable to either of these two because the midpoint rule to which it leads does not pick out the future or the past as a preferred direction in time. The midpoint rule was also originally advocated by Feynman {equation (20) in [4]} as the natural skeletonization of a path.

[40], which has the property that

$$\begin{aligned}\langle \mathbf{A}''_\perp | \widehat{\mathcal{W}} \left(F \left[\mathbf{A}_\perp, \boldsymbol{\pi}_\perp \right] \right) | \mathbf{A}'_\perp \rangle &= \int \mathcal{D}^2 \pi'_\perp F \left[\frac{\mathbf{A}'_\perp + \mathbf{A}''_\perp}{2}, \boldsymbol{\pi}'_\perp \right] \langle \mathbf{A}''_\perp | \boldsymbol{\pi}'_\perp \rangle \langle \boldsymbol{\pi}'_\perp | \mathbf{A}'_\perp \rangle \\ &= \int \mathcal{D}^2 \pi'_\perp F \left[\frac{\mathbf{A}'_\perp + \mathbf{A}''_\perp}{2}, \boldsymbol{\pi}'_\perp \right] \exp \left[i \int d^3x (\mathbf{A}''_\perp - \mathbf{A}'_\perp) \cdot \boldsymbol{\pi}'_\perp \right]. \end{aligned} \tag{III.3.29}$$

Thus, defining $\widehat{\mathcal{H}}_{\mathrm{red}} = \widehat{\mathcal{W}}(\mathcal{H}_{\mathrm{red}})$, we can rewrite

$$\begin{aligned} &\int \mathcal{D}^2 \pi^M_\perp \exp \left\{ i\delta t \int d^3x \left(\dot{\mathbf{A}}^M_\perp \cdot \boldsymbol{\pi}^M_\perp - \mathcal{H}_{\mathrm{red}} \left[\overline{\mathbf{A}}^M_\perp, \boldsymbol{\pi}^M_\perp \right] \right) \right\} \\ &\qquad = \int \mathcal{D}^2 \pi^M_\perp \left\{ 1 - i\delta t \int d^3x \mathcal{H}_{\mathrm{red}} \left[\frac{\mathbf{A}^M_\perp + \mathbf{A}^{M+1}_\perp}{2}, \boldsymbol{\pi}^M_\perp \right] + \mathcal{O}\big((\delta t)^2 \big) \right\} \\ &\qquad\qquad \times \exp \left[i \int d^3x \left(\mathbf{A}^{M+1}_\perp - \mathbf{A}^M_\perp \right) \cdot \boldsymbol{\pi}^M_\perp \right] \\ &\qquad = \left\langle \mathbf{A}^{M+1}_\perp \right| \left[1 - i\delta t \int d^3x \widehat{\mathcal{H}}_{\mathrm{red}} + \mathcal{O}\big((\delta t)^2 \big) \right] \left| \mathbf{A}^M_\perp \right\rangle , \end{aligned} \tag{III.3.30}$$

so that, dropping terms of order $(\delta t)^2$,

$$\begin{aligned} \langle \Phi | C_u | \Psi \rangle = &\int \mathcal{D}^2 A^{J+1}_\perp \left\langle \Phi(t'') \left| \mathbf{A}^{J+1}_\perp \right. \right\rangle \det[\partial_t] \det[-\partial_n] \delta T^{-1} \\ &\times \left[\prod_{M=J}^{0} \int \mathcal{D}^2 A^M_\perp \left\langle \mathbf{A}^{M+1}_\perp \right| \exp \left(-i\delta t \int d^3x \widehat{\mathcal{H}}_{\mathrm{red}} \right) \left| \mathbf{A}^M_\perp \right\rangle \right] \left\langle \mathbf{A}^0_\perp \right| \Psi(t') \rangle . \end{aligned} \tag{III.3.31}$$

Using the fact that

$$\int \mathcal{D}^2 A^M_\perp | \mathbf{A}^M_\perp \rangle \langle \mathbf{A}^M_\perp | = 1, \tag{III.3.32}$$

we see that this differs from the operator expression

$$\langle \Phi | e^{-i\widehat{H}_{\mathrm{red}} T} | \Psi \rangle \tag{III.3.33}$$

only by the factor of $\det[\partial_t] \det[-\partial_n] \delta T^{-1}$, which is a constant. From (I.3.12), we see that multiplying the class operator by a constant factor has no effect on the decoherence functional.

III.3.5 Gauge independence

We will now show that the theory described by (III.3.6) is independent of the choice of gauge G. We do this explicitly and in detail because the standard demonstration [25] makes use of a canonical transformation, which should be ill-defined at the endpoints of the integration due to the fact that the path integral (III.3.6) has one more configuration space integration than phase space integration. Put otherwise, if $\overline{\mathbf{A}}^M$ and $\boldsymbol{\pi}^M$ are linked by a canonical transformation, there is no $\boldsymbol{\pi}^{J+1}$ to which the extra degree of freedom $\overline{\mathbf{A}}^{J+1}$ corresponds.

First, we must describe how to implement an infinitesimal gauge transformation

$$\delta A^a_\mu = -\nabla_\mu \delta\Lambda^a - g f^c_{ab} A^c_\mu \delta\Lambda^b \tag{III.3.34a}$$

$$\delta\boldsymbol{\pi}^a = -g f^c_{ab} \boldsymbol{\pi}^c \delta\Lambda^b \tag{III.3.34b}$$

on a lattice. We replace the continuous function $\delta\Lambda(x)$ with functions $\left\{\delta\Lambda^M(\mathbf{x})\right\}$ defined on lattice slices 0 through $J+1$. The transformation for the vector potential can then be taken as

$$\delta\mathbf{A}^M_a = -\boldsymbol{\nabla}\delta\Lambda^M_a - g f^c_{ab}\mathbf{A}^M_c \delta\Lambda^M_b, \tag{III.3.35}$$

but there is no simple translation of the scalar potential transformation law because of the time derivative. More practical than the transformation of $\mathbf{A}^M$ given in (III.3.35) will be its implications for transformations of the time derivative $\dot{\mathbf{A}}^M$ and midpoint average $\overline{\mathbf{A}}^M$. Using the fact that

$$\begin{aligned}\frac{A^{M+1}B^{M+1} - A^M B^M}{\delta t} &= \frac{\left(A^{M+1} - A^M\right)}{\delta t}\frac{\left(B^{M+1} + B^M\right)}{2} \\ &\quad + \frac{\left(A^{M+1} + A^M\right)}{2}\frac{\left(B^{M+1} - B^M\right)}{\delta t} = \dot{A}^M \overline{B}^M + \overline{A}^M \dot{B}^M,\end{aligned} \tag{III.3.36}$$

we see that (III.3.35) implies

$$\delta\dot{\mathbf{A}}^M_a = -\boldsymbol{\nabla}\delta\dot{\Lambda}^M_a - g f^c_{ab}\left(\dot{\mathbf{A}}^M_c \overline{\delta\Lambda}^M_b + \overline{\mathbf{A}}^M_c \delta\dot{\Lambda}^M_b\right), \tag{III.3.37}$$

which is exactly what one might write down from the corresponding continuum expression

$$\delta\dot{\mathbf{A}}_a = -\boldsymbol{\nabla}\delta\dot{\Lambda}_a - g f^c_{ab}\left(\dot{\mathbf{A}}_c \delta\Lambda_b + \mathbf{A}_c \delta\dot{\Lambda}_b\right). \tag{III.3.38}$$

The case is not quite so simple with the averages. Since

$$\begin{aligned}\frac{1}{2}\left(A^{M+1}B^{M+1} + A^M B^M\right) &= \frac{1}{4}\left(A^{M+1} + A^M\right)\left(B^{M+1} + B^M\right) \\ &\quad + \frac{1}{4}\left(A^{M+1} - A^M\right)\left(B^{M+1} - B^M\right) = \overline{A}^M \overline{B}^M + \frac{(\delta t)^2}{4}\dot{A}^M \dot{B}^M,\end{aligned} \tag{III.3.39}$$

we have

$$\delta\overline{\mathbf{A}}^M_a = -\boldsymbol{\nabla}\overline{\delta\Lambda}^M_a - g f^c_{ab}\left(\overline{\mathbf{A}}^M_c \overline{\delta\Lambda}^M_b + \frac{(\delta t)^2}{4}\dot{\mathbf{A}}^M_c \delta\dot{\Lambda}^M_b\right), \tag{III.3.40}$$

which differs from the naïve analog of (III.3.35) by a term proportional to $(\delta t)^2$. However, we can neglect this term by the following familiar argument [41]: as the lattice spacing δt goes to zero, the factor of $\exp[-i\delta t \int d^3x(\boldsymbol{\pi}^2/2 - \boldsymbol{\pi}\cdot\dot{\mathbf{A}})]$ will oscillate rapidly and suppress the path integral if $\dot{\mathbf{A}}$ is more singular than $(\delta t)^{-1/2}$. Likewise, if we concern ourselves only with gauge transformations which take histories which are sufficiently nonsingular to contribute to the path integral into other such histories, (III.3.37) tells us that the $\delta\dot{\Lambda}$ should also blow up no faster than $(\delta t)^{-1/2}$ as $\delta t \to 0$. Counting the factors of δt in the upper bounds, we see that the extra term in (III.3.40) should be at worst proportional to δt and thus be negligible for sufficiently small lattice spacing. Thus we can use the simpler formula

$$\delta\overline{\mathbf{A}}_a^M = -\boldsymbol{\nabla}\overline{\delta\Lambda}_a^M - g f_{ab}^c \overline{\mathbf{A}}_c^M \overline{\delta\Lambda}_b^M. \tag{III.3.40$'$}$$

The transformation laws for the conjugate momentum and scalar potential can be defined by analogy to (III.3.40′). As observed in Sec. III.3.4 the scalar potential only enters the class operator via its averaged values (assuming that $\overline{\varphi}^{J+1} = 0$ is always taken as a gauge choice), and so we only need to know how to transform $\overline{\varphi}$ and not φ. Since $\overline{\varphi}$ and $\boldsymbol{\pi}$ are defined midway between lattice points, we prescribe transformation laws which are the obvious lattice realizations of (III.3.34):

$$\delta\overline{\varphi}_a^M = \delta\dot{\Lambda}_a^M - g f_{ab}^c \overline{\varphi}_c^M \overline{\delta\Lambda}_b^M \tag{III.3.41a}$$

$$\delta\boldsymbol{\pi}_a^M = -g f_{ab}^c \boldsymbol{\pi}_c^M \overline{\delta\Lambda}_b^M. \tag{III.3.41b}$$

Now we show that under such a gauge transformation, the expression (III.3.6) for the class operator is unchanged. First, we examine the measure for the path integral. The demonstrations for $\mathcal{D}^3 A^M$, $\mathcal{D}\overline{\varphi}^M$ and $\mathcal{D}^3\pi^M$ are all essentially the same, so we show it explicitly only for $\mathcal{D}^3 A^M$. Under the gauge transformation $\mathbf{A}^M \to \widetilde{\mathbf{A}}^M = \mathbf{A}^M + \delta\mathbf{A}^M$, we have

$$\prod_a d\widetilde{A}_a^{Mi}(\mathbf{x}) = \left(\prod_a dA_a^{Mi}(\mathbf{x})\right) \det\left\{\frac{\partial \widetilde{A}_b^{Mi}(\mathbf{x})}{\partial A_c^{Mi}(\mathbf{x})}\right\}. \tag{III.3.42}$$

The Jacobian matrix is

$$\frac{\partial \widetilde{A}_b^{Mi}(\mathbf{x})}{\partial A_c^{Mi}(\mathbf{x})} = \delta_{bc} + \frac{\partial\, \delta A_b^{Mi}(\mathbf{x})}{\partial A_c^{Mi}(\mathbf{x})} = \delta_{bc} - g f_{bd}^c \delta\Lambda_d^M(\mathbf{x}). \tag{III.3.43}$$

Using the standard matrix result that to lowest order in δa, $\det(1+\delta a) = 1 + \operatorname{Tr}\delta a$, we see that the Jacobian for the transformation is

$$\det\left\{\frac{\partial \widetilde{A}_b^{Mi}(\mathbf{x})}{\partial A_c^{Mi}(\mathbf{x})}\right\} = 1 - f_{bd}^b \delta\Lambda_d^M(\mathbf{x}) = 1 \tag{III.3.44}$$

where we have used the fact that the structure constants are totally antisymmetric. This tells us that

$$\mathcal{D}^3\widetilde{A}^M = \mathcal{D}^3 A^M, \tag{III.3.45a}$$

and the demonstrations that

$$\mathcal{D}\widetilde{\overline{\varphi}}^M = \mathcal{D}\overline{\varphi}^M \tag{III.3.45b}$$

and

$$\mathcal{D}^3\widetilde{\pi}^M = \mathcal{D}^3\pi^M \tag{III.3.45c}$$

proceed similarly.

Next we consider the canonical action density [i.e., $\dot{\mathbf{A}} \cdot \boldsymbol{\pi}$ minus the Hamiltonian density; *cf.* (III.3.3)]. The demonstration is simplest if we undo the integration by parts to write it in the form

$$\dot{\mathbf{A}}^M \cdot \boldsymbol{\pi}^M + \boldsymbol{\pi}^M \cdot \overline{\mathbf{D}}^M \overline{\varphi}^M - \frac{1}{2}\left(\boldsymbol{\pi}^M\right)^2 - \frac{1}{2}\left(\mathbf{B}^M\right)^2 = -\mathbf{E}^M \cdot \boldsymbol{\pi}^M - \frac{1}{2}\left(\boldsymbol{\pi}^M\right)^2 - \frac{1}{2}\left(\mathbf{B}^M\right)^2, \quad \text{(III.3.46)}$$

where we have defined the lattice realizations of the electric and magnetic fields by analogy to (III.2.11) and (III.2.8b):

$$\mathbf{E}^M = -\dot{\mathbf{A}}^M - \overline{\mathbf{D}}^M \overline{\varphi}^M \qquad \text{(III.3.47a)}$$

and

$$\mathbf{B}_a^M = \boldsymbol{\nabla} \times \overline{\mathbf{A}}_a^M + \frac{1}{2} g f_{ab}^c \overline{\mathbf{A}}_c^M \times \overline{\mathbf{A}}_b^M . \qquad \text{(III.3.47b)}$$

The conjugate momentum $\boldsymbol{\pi}$ is defined by (III.3.41b) to transform as an isovector. The transformations of $\mathbf{E}$ and $\mathbf{B}$ are

$$\delta\mathbf{E}_a^M = -\delta\dot{\mathbf{A}}_a^M - \left(\mathbf{D}\delta\overline{\varphi}^M\right)_a - g f_{ab}^c \delta\overline{\mathbf{A}}_c^M \overline{\varphi}_b^M \qquad \text{(III.3.48a)}$$

$$\delta\mathbf{B}_a^M = \boldsymbol{\nabla} \times \delta\overline{\mathbf{A}}_a^M + g f_{ab}^c \delta\overline{\mathbf{A}}_c^M \times \overline{\mathbf{A}}_b^M , \qquad \text{(III.3.48b)}$$

which can be shown, with a little algebra, to give the expected isovector transformations:

$$\delta\mathbf{E}_a^M = -g f_{ab}^c \mathbf{E}_c^M \overline{\delta\Lambda}_b^M \qquad \text{(III.3.49a)}$$

$$\delta\mathbf{B}_a^M = -g f_{ab}^c \mathbf{B}_c^M \overline{\delta\Lambda}_b^M . \qquad \text{(III.3.49b)}$$

So, since $\boldsymbol{\pi}^M$, $\mathbf{E}^M$ and $\mathbf{B}^M$ all transform as isovectors, $\left(\boldsymbol{\pi}^M\right)^2$, $\left(\mathbf{B}^M\right)^2$ and $\mathbf{E}^M \cdot \boldsymbol{\pi}^M$ are gauge invariant quantities, and

$$\begin{aligned} &\dot{\widetilde{\mathbf{A}}}^M \cdot \widetilde{\boldsymbol{\pi}}^M + \widetilde{\boldsymbol{\pi}}^M \cdot \overline{\widetilde{\mathbf{D}}}^M \overline{\widetilde{\varphi}}^M - \frac{1}{2}\left(\widetilde{\boldsymbol{\pi}}^M\right)^2 - \frac{1}{2}\left(\widetilde{\mathbf{B}}^M\right)^2 \\ &\qquad = \dot{\mathbf{A}}^M \cdot \boldsymbol{\pi}^M + \boldsymbol{\pi}^M \cdot \overline{\mathbf{D}}^M \overline{\varphi}^M - \frac{1}{2}\left(\boldsymbol{\pi}^M\right)^2 - \frac{1}{2}\left(\mathbf{B}^M\right)^2 . \end{aligned} \qquad \text{(III.3.50)}$$

The gauge-fixing delta function becomes $\delta\left[\widetilde{G}^M\right]$. Using (III.3.2c) to write the Fadeev-Popov determinant as

$$\frac{1}{\Delta_G} = \int \mathcal{D}\Lambda' \delta[G^{\Lambda'} - G], \qquad \text{(III.3.51)}$$

where $G^{\Lambda'}$ indicates the result of a gauge transformation on G by the dummy variable $\Lambda'(x)$, we see that the effect of a gauge transformation by $\Lambda(x)$ is

$$\frac{1}{\widetilde{\Delta}_G} = \frac{1}{\Delta_G^{\Lambda}} = \int \mathcal{D}\Lambda' \delta[G^{\Lambda' \cdot \Lambda} - G^{\Lambda}]. \qquad \text{(III.3.52)}$$

Changing the variable of integration to $\Lambda'' = \Lambda' \cdot \Lambda$ (the gauge transformation accomplished by successive application of Λ' and Λ), we have (using formal invariance of the group measure; see section 7.5 of [39] for more details)

$$\frac{1}{\widetilde{\Delta}_G} = \int \mathcal{D}\Lambda'' \delta[G^{\Lambda''} - \widetilde{G}] = \frac{1}{\Delta_{\widetilde{G}}}. \tag{III.3.53}$$

Coarse graining only by gauge-invariant alternatives means that a path $\{\widetilde{A}, \widetilde{\boldsymbol{\pi}}\}$ is in the class c_α if and only if the corresponding path $\{A, \boldsymbol{\pi}\}$ is, so

$$e_\alpha[\widetilde{A}, \widetilde{\boldsymbol{\pi}}] = e_\alpha[A, \boldsymbol{\pi}]. \tag{III.3.54}$$

Finally, we consider the behavior of the wave functionals Φ^* and Ψ.

$$\begin{aligned} \Psi[\widetilde{\mathbf{A}}'] &= \Psi[\mathbf{A}' + \delta\mathbf{A}'] = \Psi[\mathbf{A}'] + \int d^3x \, \delta\mathbf{A}'(\mathbf{x}) \cdot \frac{\mathcal{D}\Psi}{\mathcal{D}\mathbf{A}'(\mathbf{x})} \\ &= \Psi[\mathbf{A}'] - \int d^3x \, (\mathbf{D}'\delta\Lambda')(\mathbf{x}) \cdot \frac{\mathcal{D}\Psi}{\mathcal{D}\mathbf{A}'(\mathbf{x})}. \end{aligned} \tag{III.3.55}$$

Upon integrating by parts, this becomes

$$\Psi[\widetilde{\mathbf{A}}'] = \Psi[\mathbf{A}'] + \int d^3x \, \delta\Lambda' \, \mathbf{D}' \cdot \frac{\mathcal{D}\Psi}{\mathcal{D}\mathbf{A}'} = \Psi[\mathbf{A}'], \tag{III.3.56}$$

where we have used the operator constraint (III.2.18b) on the wave functional Ψ in the last step.

Now, we relabel the variables A and $\boldsymbol{\pi}$ in (III.3.6) by $\widetilde{A}$ and $\widetilde{\boldsymbol{\pi}}$ and use (III.3.45), (III.3.50), (III.3.53), (III.3.54) and (III.3.56) to convert the expression to

$$\begin{aligned} \langle \Phi | C_\alpha | \Psi \rangle &= \int \mathcal{D}^3 \widetilde{A}^{J+1} \Phi^* \left[\widetilde{\mathbf{A}}^{J+1}; t'' \right) \\ &\times \left(\prod_{M=J}^{0} \int \mathcal{D}^3 \widetilde{A}^M \mathcal{D} \widetilde{\overline{\varphi}}^M \mathcal{D}^3 \widetilde{\pi}^M \exp \left\{ i\delta t \int d^3x \left(\dot{\widetilde{\mathbf{A}}}^M \cdot \widetilde{\boldsymbol{\pi}}^M - \mathcal{H}\left[\widetilde{\overline{\mathbf{A}}}^M, \widetilde{\boldsymbol{\pi}}^M \right] - \widetilde{\overline{\varphi}}^M \widetilde{\overline{\mathbf{D}}}^M \cdot \widetilde{\boldsymbol{\pi}}^M \right) \right\} \right) \\ &\qquad \times \left(\prod_{M=0}^{J+1} \delta\left[\widetilde{G}^M \right] \widetilde{\Delta}_{G^M} \right) \Psi\left[\widetilde{\mathbf{A}}^0; t' \right) e_\alpha \left[\widetilde{A}, \widetilde{\boldsymbol{\pi}} \right] \\ &= \int \mathcal{D}^3 A^{J+1} \Phi^* \left[\mathbf{A}^{J+1}; t'' \right) \\ &\times \left(\prod_{M=J}^{0} \int \mathcal{D}^3 A^M \mathcal{D} \overline{\varphi}^M \mathcal{D}^3 \pi^M \exp \left\{ i\delta t \int d^3x \left(\dot{\mathbf{A}}^M \cdot \boldsymbol{\pi}^M - \mathcal{H}\left[\overline{\mathbf{A}}^M, \boldsymbol{\pi}^M \right] - \overline{\varphi}^M \overline{\mathbf{D}}^M \cdot \boldsymbol{\pi}^M \right) \right\} \right) \\ &\qquad \times \left(\prod_{M=0}^{J+1} \delta\left[\widetilde{G}^M \right] \Delta_{\widetilde{G}^M} \right) \Psi\left[\mathbf{A}^0; t' \right) e_\alpha [A, \boldsymbol{\pi}], \end{aligned} \tag{III.3.57}$$

which shows that the expressions for the class operators are the same whether the gauge is $G = 0$ or $\widetilde{G} = 0$.

III.4 Phase space results

III.4.1 Allowed alternatives

In the previous section, we described a sum-over-histories construction of the class operators (and hence the decoherence functional) for a NAGT. For the class operators to have the desirable properties detailed therein, the alternatives considered need only be gauge-invariant coarse grainings of the connection A and the conjugate momentum $\boldsymbol{\pi}$. However, in practice we will not be interested in arbitrary gauge-invariant quantities, but only the physical quantities of the theory, namely the gauge electric and magnetic fields and the covariant derivative. In a phase space formulation, those are identified with the isovectors $-\boldsymbol{\pi}$, $\mathbf{B}$ and D_μ, respectively. Thus we define physical phase space coarse grainings to be those in which the gauge electric field is identified with $-\boldsymbol{\pi}$ and the gauge magnetic field with $\mathbf{B}$. (The physical phase space theory is then a subset of the phase space theory described in section III.3.)

We can partition the histories by the values of arbitrary isoscalars (gauge invariant quantities) constructed from the physically allowed isovectors. Since the length of an isovector is an isoscalar, coarse graining by the length of isovectors is allowed. It is more convenient to think of this sort of coarse graining as specifying in which of a set of regions in isospace an isovector lies, where all the regions are rotationally invariant. From now on, when we talk about coarse graining by isovectors, this is what we mean.

III.4.2 Constraints

So a general "physical" coarse graining can involve functionals of $\boldsymbol{\pi}$, $\mathbf{B}$, $\mathbf{D}$, and D_t. If we consider the subset of coarse grainings which involves the first three but not the covariant time derivative D_t, we see that it involves only the vector potential $\mathbf{A}$ and the conjugate momentum $\boldsymbol{\pi}$, and not the scalar potential φ. If we work in a gauge which also leaves φ unrestricted, we can perform the path integral over φ to obtain

$$\int \mathcal{D}\varphi \exp\left(i \int d^4x\, \varphi\, \mathbf{D}\cdot\boldsymbol{\pi}\right) = \delta[\mathbf{D}\cdot\boldsymbol{\pi}/2\pi], \tag{III.4.1}$$

and (III.3.1) becomes

$$\langle\Phi|C_\alpha|\Psi\rangle = \int_\alpha \mathcal{D}^3A\mathcal{D}^3\pi\Phi^*[\mathbf{A}'';t'')\det[2\pi]\delta[K]\delta[G]\Delta_G[A,\boldsymbol{\pi}]e^{iS_{\mathrm{can}}[A,\boldsymbol{\pi}]}\Psi[\mathbf{A}';t'). \tag{III.4.2}$$

In particular, if we coarse grain by values of the constraint $K = \mathbf{D}\cdot\boldsymbol{\pi}$, the delta function will cause the class operator to vanish for any class which does not include $\mathbf{D}\cdot\boldsymbol{\pi} = 0$, i.e., the constraint satisfied. If exactly one class includes that condition, then, it will have the only non-vanishing class operator, and the only non-vanishing element of the decoherence functional will be the diagonal element corresponding to that alternative. This will then allow the assignment of probabilities, namely, a probability of 1 for the alternative in which the constraint is satisfied and 0 for all others.

Now we make this formal demonstration more precise. Given our choice of lattice expressions, it should be clear that the relevant quantities are the following functions defined on each slice:

$$\left\{K^M(\mathbf{x}) = \left(\overline{\mathbf{D}}^M \cdot \boldsymbol{\pi}^M\right)(\mathbf{x})\right\}. \qquad \text{(III.4.3)}$$

If we work in a gauge which sets $\overline{\varphi}^{J+1}$ to zero and does not otherwise restrict φ, we can perform the $\overline{\varphi}$ integrals as in (III.3.17) and obtain from (III.3.6) [recalling (III.3.12)]

$$\begin{aligned}\langle\Phi|C_\alpha|\Psi\rangle = &\int \mathcal{D}^3A^{J+1}\Phi^*\left[\mathbf{A}^{J+1};t''\right)\det[\partial_t] \\ &\times\left(\prod_{M=J}^{0}\int \mathcal{D}^3A^M\mathcal{D}^3\pi^M \exp\left\{i\delta t\int d^3x\left(\dot{\mathbf{A}}^M\cdot\boldsymbol{\pi}^M - \mathcal{H}\left[\overline{\mathbf{A}}^M,\boldsymbol{\pi}^M\right]\right)\right\}\delta\left[K^M\right]\delta T^{-1}\right) \\ &\times\left(\prod_{M=0}^{J+1}\delta\left[G^M\right]\Delta_{G^M}\right)\Psi\left[\mathbf{A}^0;t'\right)e_\alpha\left[A,\boldsymbol{\pi}\right].\end{aligned} \qquad \text{(III.4.4)}$$

The $\delta\left[K^M\right]$ causes the class operator to vanish for any coarse graining not consistent with all the $\left\{K^M\right\}$ vanishing everywhere. Thus if we coarse grain by an average of the constraint over some spacetime region, which will correspond to some average over the $\left\{K^M\right\}$, the only non-zero element of the decoherence functional will be the diagonal one corresponding to that average vanishing.

III.4.3 Comparison to reduced phase space theory

In a reduced phase space canonical theory, as described in section III.3.4, the alternatives are defined by projections onto eigenstates of physical operators. Working in the axial gauge, the operator corresponding to the gauge electric field is [*cf.* (III.3.19)]

$$-\left(\widehat{\boldsymbol{\pi}}_\perp - \mathbf{e}_n\partial_n^{-1}\widehat{\mathbf{D}}_\perp\cdot\widehat{\boldsymbol{\pi}}_\perp\right) \qquad \text{(III.4.5a)}$$

and that corresponding to the gauge magnetic field is

$$\widehat{\mathbf{B}}^a = \boldsymbol{\nabla}\times\widehat{\mathbf{A}}^a_\perp + \frac{1}{2}gf^c_{ab}\widehat{\mathbf{A}}^c_\perp\times\widehat{\mathbf{A}}^b_\perp. \qquad \text{(III.4.5b)}$$

The operator for the covariant gradient is $\widehat{\mathbf{D}}^{ab} = \delta_{ab}\boldsymbol{\nabla} + gf^c_{ab}\widehat{\mathbf{A}}^c_\perp$, but it is less clear how to convert $D^{ab}_t = \delta_{ab}\partial_t - gf^c_{ab}\varphi_c$ into an operator built out of $\widehat{\mathbf{A}}_\perp$ and $\widehat{\boldsymbol{\pi}}_\perp$. First of all, we need to consider the time derivatives of $\widehat{\mathbf{A}}_\perp$ and $\widehat{\boldsymbol{\pi}}_\perp$. In an operator theory, one identifies the time derivatives at one moment of time via the Heisenberg equations of motion:

$$\partial_t\widehat{\mathbf{A}}_\perp = i\left[\widehat{H}_{\text{red}},\widehat{\mathbf{A}}_\perp\right] = \widehat{\boldsymbol{\pi}}_\perp + \widehat{\mathcal{W}}\left(\mathbf{D}_\perp\partial_n^{-2}\mathbf{D}_\perp\cdot\boldsymbol{\pi}_\perp\right) \qquad \text{(III.4.6a)}$$

$$\partial_t\widehat{\boldsymbol{\pi}}^a_\perp = i\left[\widehat{H}_{\text{red}},\widehat{\boldsymbol{\pi}}^a_\perp\right] = -\mathbb{I}_\perp\cdot\left(\widehat{\mathbf{D}}_\perp\times\widehat{\mathbf{B}}\right)^a - gf^c_{ab}\widehat{\mathcal{W}}\Big(\left(\partial_n^{-2}\mathbf{D}_\perp\cdot\boldsymbol{\pi}_\perp\right)^c\boldsymbol{\pi}^b_\perp\Big), \qquad \text{(III.4.6b)}$$

where

$$\mathbb{I}_\perp = \mathbb{I} - \mathbf{e}_n \otimes \mathbf{e}_n \tag{III.4.7}$$

is the tensor which projects onto the "perpendicular" directions. If we identify

$$\widehat{\varphi} = -\partial_n^{-2} \widehat{\mathbf{D}}_\perp \cdot \widehat{\boldsymbol{\pi}}_\perp, \tag{III.4.8}$$

and remember that the final operator expressions should always be Weyl-ordered, equations (III.4.6) give the operator versions of $-\mathbf{E}_\perp = \boldsymbol{\pi}_\perp$ [two of the components of (III.2.16a)] and the Maxwell's equation $D_t \boldsymbol{\pi}_\perp = -(\mathbf{D} \times \mathbf{B})_\perp$ [two of the components of (III.2.16b)], respectively. We also have

$$\widehat{\pi}_n = \partial_n \widehat{\varphi}, \tag{III.4.9}$$

which is the third component of (III.2.16a). So the operator realization of D_t^{ab} is

$$\delta_{ab} \partial_t + g f_{ab}^c \partial_n^{-2} \left(\widehat{\mathbf{D}}_\perp \cdot \widehat{\boldsymbol{\pi}}_\perp \right)_c, \tag{III.4.10}$$

where the effects of ∂_t on other operators are given by (III.4.6). Of course, in a sum-over-histories formulation, attempts to describe instantaneous values of time derivatives do not yield sensible results due to the non-differentiability of the paths. Instead, we coarse grain by time derivatives averaged over time, which correspond to coarse grainings by the difference between values of a quantity at two finitely separated instants of time.

As described in section I.3.3, probabilities in the usual operator quantum mechanics are described by expectation values of projection operators, while in an operator generalized quantum mechanics [5], an alternative c_α corresponds to a class operator C_α which is defined as in (I.3.8) by a chain of such projections:

$$C_\alpha = e^{-i\widehat{H}_{\mathrm{red}}(t''-t_n)} P^n_{\alpha_n} \left(\prod_{i=n-1}^{1} e^{-i\widehat{H}_{\mathrm{red}}(t_{i+1}-t_i)} P^i_{\alpha_i} \right) e^{-i\widehat{H}_{\mathrm{red}}(t_1-t')}. \tag{III.4.11}$$

If we generalize further, and allow a class operator to be not just a single chain of projections, but a sum of such chains, we can describe more general alternatives, such as coarse grainings by time averages. The operator expression corresponding to a coarse-grained class operator would be defined by assigning to each class c_α a sum of chains of projections:

$$C_\alpha = \sum_{\{\alpha_i\}\in\alpha} e^{-i\widehat{H}_{\mathrm{red}}(t''-t_n)} P^n_{\alpha_n} \left(\prod_{i=n-1}^{1} e^{-i\widehat{H}_{\mathrm{red}}(t_{i+1}-t_i)} P^i_{\alpha_i} \right) e^{-i\widehat{H}_{\mathrm{red}}(t_1-t')}. \tag{III.4.11$'$}$$

This will be equivalent to the corresponding sum-over-histories expression if we can replace the projections with restricted integrations on lattice slices. First, consider a projection onto a range of the coördinate $\mathbf{A}_\perp$:

$$P^i_{\Delta_i} = \int \mathcal{D}^2 A^\iota_\perp \left| \mathbf{A}^\iota_\perp \right\rangle e_{\Delta_i}[\mathbf{A}^\iota_\perp] \left\langle \mathbf{A}^\iota_\perp \right|, \tag{III.4.12}$$

where

$$e_{\Delta_i}[\mathbf{A}^\iota_\perp] = \begin{cases} 0, & \mathbf{A}^\iota_\perp \notin \Delta_i \\ 1, & \mathbf{A}^\iota_\perp \in \Delta_i \end{cases} \tag{III.4.13}$$

is the indicator function for the region Δ_i in the space of field configurations $\{\mathbf{A}^\iota_\perp(\mathbf{x})\}$. To examine the effect that this projection has on the class operator, assume that we have taken our lattice spacing small enough that $\delta t < t_{i+1} - t_i, t_i - t_{i-1}$ so that if t_I is the latest time slice before t_i (i.e., $t_{I+1} \geq t_i \geq t_I$), no other projections lie in the interval (t_I, t_{I+1}). Then the effect of the projection is to modify the right-hand side of (III.3.30) to be, to first order in δt,

$$\begin{aligned} &\left\langle \mathbf{A}_\perp^{I+1} \right| \left[1 - i(t_{I+1} - t_i)\widehat{H}_{\mathrm{red}}\right] P^i_{\Delta i} \left[1 - i(t_i - t_I)\widehat{H}_{\mathrm{red}}\right] \left| \mathbf{A}^I_\perp \right\rangle \\ &\quad = \left\langle \mathbf{A}_\perp^{I+1} \right| \left(1 - i\delta t \widehat{H}_{\mathrm{red}}\right) \left| \mathbf{A}^I_\perp \right\rangle \left(\frac{t_{I+1} - t_i}{\delta t} e_{\Delta_i}\left[A_\perp^{I+1}\right] + \frac{t_i - t_I}{\delta t} e_{\Delta_i}\left[A^I_\perp\right] \right). \end{aligned} \tag{III.4.14}$$

In the sum-over-histories formulation, we would choose an $\mathbf{A}^\iota_\perp$ most closely corresponding to $\mathbf{A}_\perp(t_i)$ and one factor in the indicator function e_α would be $e_{\Delta_i}[\mathbf{A}^\iota_\perp]$. Given the spirit of our lattice resolution, that is clearly $\overline{\mathbf{A}}^I_\perp$. There is a discrepancy between the expressions

$$e_{\Delta_i}\left[\overline{\mathbf{A}}^I_\perp\right] \tag{III.4.15a}$$

and

$$\frac{t_{I+1} - t_i}{\delta t} e_{\Delta_i}\left[\mathbf{A}_\perp^{I+1}\right] + \frac{t_i - t_I}{\delta t} e_{\Delta_i}\left[\mathbf{A}^I_\perp\right], \tag{III.4.15b}$$

even when $t_i = t_I + \frac{\delta t}{2}$. However, it can be seen as an artifact of the lattice; we argued in section III.3.5 that as $\delta t \to 0$, the difference between $\mathbf{A}^I_\perp$, $\mathbf{A}^{I+1}_\perp$, and $\overline{\mathbf{A}}^I_\perp$ vanishes like $(\delta t)^{1/2}$, so we expect both expressions to give the same results in that limit. {It is interesting to note that (III.4.15a) and (III.4.15b) are reminiscent of equations (20) and (19), respectively, of Feynman's original paper [4] on path integrals.}

Here we should consider a moment just what "projections onto ranges of $\mathbf{A}_\perp$" means, physically. After all, $\mathbf{A}$ is a gauge-dependent quantity, so it cannot be the expression which determines the coarse graining in (III.3.6) independent of the gauge choice G. The two quantities of interest constructed from $\mathbf{A}$ are $\mathbf{D}$ and $\mathbf{B}$. Since $\mathbf{B}$, as defined by (III.2.8b), depends only on the vector potential $\mathbf{A}$ at a single time, a field configuration $\mathbf{B}(\mathbf{x})$ can be determined from a field configuration $\mathbf{A}(\mathbf{x})$ alone. In fact, the gauge freedom means there are many configurations of $\mathbf{A}$ which lead to the same $\mathbf{B}$ configuration, so there is a one-to-one correspondence between field configurations $\mathbf{B}(\mathbf{x})$ and gauge-fixed field configurations $\mathbf{A}_\perp(\mathbf{x})$.

If we project by values of $\boldsymbol{\pi}$,

$$P^i_{\Delta_i} = \int \mathcal{D}^2\pi^\iota_\perp \left|\boldsymbol{\pi}^\iota_\perp\right\rangle e_{\Delta_i}[\boldsymbol{\pi}^\iota_\perp] \left\langle\boldsymbol{\pi}^\iota_\perp\right|, \tag{III.4.16}$$

we find that to first order in δt

$$\begin{aligned}\left\langle \mathbf{A}_\perp^{I+1}\right| e^{-i\widehat{H}_{\rm red}(t_{I+1}-t_i)} P^i_{\Delta_i} e^{-i\widehat{H}_{\rm red}(t_i-t_I)} \left|\mathbf{A}_\perp^{I}\right\rangle \\ = \int \mathcal{D}^2\pi_\perp^\kappa \mathcal{D}^2 A_\perp^\iota \mathcal{D}^2\pi_\perp^\iota \left\langle \mathbf{A}_\perp^{I+1}\right| \boldsymbol{\pi}_\perp^\kappa\rangle \langle \boldsymbol{\pi}_\perp^\kappa \left|\mathbf{A}_\perp^{\iota}\right\rangle \left\langle \mathbf{A}_\perp^{\iota}\right| \boldsymbol{\pi}_\perp^\iota\rangle \langle \boldsymbol{\pi}_\perp^\iota \left|\mathbf{A}_\perp^{I}\right\rangle \\ \times \left\{ \frac{t_{I+1}-t_i}{\delta t} \left(1 - i\delta t H_{\rm red}\left[\frac{\mathbf{A}_\perp^{I+1}+\mathbf{A}_\perp^\iota}{2}, \boldsymbol{\pi}_\perp^\kappa\right]\right) e_{\Delta_i}[\boldsymbol{\pi}_\perp^\iota] \right. \\ \left. + \frac{t_i - t_I}{\delta t}\left(1 - i\delta t H_{\rm red}\left[\frac{\mathbf{A}_\perp^{\iota}+\mathbf{A}_\perp^I}{2}, \boldsymbol{\pi}_\perp^\iota\right]\right) e_{\Delta_i}[\boldsymbol{\pi}_\perp^\kappa] \right\}. \quad \text{(III.4.17)}\end{aligned}$$

If we argue that in the limit $\delta t \to 0$ we can replace $\mathbf{A}_\perp^\iota$ with $\mathbf{A}_\perp^I$ in the first term and $\mathbf{A}_\perp^{I+1}$ in the second term, we recover the sum-over-histories expression

$$\int \mathcal{D}^2\pi_\perp^\iota \left\langle \mathbf{A}_\perp^{I+1}\right| \boldsymbol{\pi}_\perp^\iota\rangle \langle \boldsymbol{\pi}_\perp^\iota \left|\mathbf{A}_\perp^{I}\right\rangle \left(1 - i\delta t H_{\rm red}\left[\frac{\mathbf{A}_\perp^{I+1}+\mathbf{A}_\perp^I}{2}, \boldsymbol{\pi}_\perp^\iota\right]\right) e_{\Delta_i}[\boldsymbol{\pi}_\perp^\iota]. \quad \text{(III.4.18)}$$

Now we consider physically what a projection onto values of $\boldsymbol{\pi}_\perp$ means. The conjugate momentum $\boldsymbol{\pi}$ is gauge-covariant, and in the physical phase space theory directly accessible, so as long as the regions $\{\Delta_i\}$ are rotationally invariant in isospace, these are allowed sets of alternatives. However, dependence of the reduced Hamiltonian (III.3.25) on $\boldsymbol{\pi}_\perp$ not just directly but through the fixed form (III.3.20) of π_n means that we are actually restricting not $\boldsymbol{\pi}_\perp$ independently, but $\boldsymbol{\pi}_\perp$ subject to the constraint $\mathbf{D}\cdot\boldsymbol{\pi} = 0$. This is no cause for alarm, though, since as long as the coarse graining makes no reference to $\varphi(t_i)$, the result (III.4.2) [or (III.4.4)] ensures that this is also the case in the sum-over-histories formulation.

It seems reasonable to assume that, modulo operator ordering delicacies, a similar correspondence will hold for any combination of $\mathbf{D}$, $\mathbf{B}$ and $\boldsymbol{\pi}$, and *the the sum-over-histories formulation gives the same results as the corresponding reduced phase space canonical theory* {referred to in [24] and elsewhere as "Arnowitt-Deser-Misner (ADM) quantization"} *for physical phase space coarse grainings not involving D_t.*

III.5 Configuration space results

III.5.1 Allowed alternatives; overview

In the sum-over-histories formulation, it is possible to consider a set of physical alternatives in which the conjugate momenta are not specified. The gauge electric field, which was previously described by $-\boldsymbol{\pi}$, is now described solely in terms of the potentials as $\mathbf{E} = -\dot{\mathbf{A}} - \mathbf{D}\varphi$. While these two definitions of the gauge electric field are classically equivalent, quantum mechanical descriptions based on them will in general be inequivalent. The physical configuration space theory is that in which the gauge electric field is represented by $\mathbf{E}$ and the gauge magnetic field by $\mathbf{B}$. It has the advantage over the physical phase space realization that, as described in section III.6, it is formally manifestly Lorentz-invariant.

However, for that very reason, it will turn out to be not completely consistent with the enforcement of the Gauss's law constraint.

In the subset of gauge-invariant alternatives which do not restrict the momenta (of which the *physical* configuration space alternatives are in turn a subset), the integrals over the momenta in the path integral (III.3.6) for the class operator can be explicitly performed. We do this in section III.5.2, which gives us a constructive definition of the configuration space path integral (III.3.5). For the purposes of the physical configuration space realization of the NAGT, we could have started with the formal expression (III.3.5), but approaching it via the phase space route has enabled us to calculate naturally the measure for the configuration space path integral.

After constructing the configuration space path integral, we spend the next two subsections of section III.5 on coarse grainings by the configuration space constraints. In section III.5.3 we construct a class of quantities which have the same properties as those found for the constraints of E&M in [24]: the only decohering coarse grainings are those which predict that the constraint is satisfied with 100% probability. In section III.5.4, however, we exhibit a quantity in the Abelian theory of E&M which vanishes in the presence of the Gauss's law constraint, yet violates this property; there are decohering coarse grainings which predict nonvanishing values of the quantity with nonzero probability.

III.5.2 Reducing the phase space formulation to the configuration space formulation

Since the coarse grainings make no reference to the canonical momenta $\boldsymbol{\pi}$, we can work in a gauge where $\boldsymbol{\pi}$ is unrestricted and perform the Gaussian integrals over the momenta in (III.3.6) [using the form of the canonical action density in (III.3.46)]:

$$\begin{aligned}
&\int \mathcal{D}^3\pi^M \exp\left\{-i\delta t \int d^3x \left[\frac{1}{2}\left(\boldsymbol{\pi}^M\right)^2 - \boldsymbol{\pi}^M\cdot\left(\dot{\mathbf{A}}^M + \overline{\mathbf{D}}^M\overline{\varphi}^M\right)\right]\right\} \\
&= \prod_{a,\mathbf{x}} N_\pi^3 \int d^3\pi_a^M(\mathbf{x}) \exp\left(-i\delta t\delta^3x\left\{\frac{1}{2}\left[\boldsymbol{\pi}_a^M(\mathbf{x})\right]^2 - \boldsymbol{\pi}_a^M(\mathbf{x})\cdot\left[\dot{\mathbf{A}}_a^M(\mathbf{x}) + \left(\overline{\mathbf{D}}^M\overline{\varphi}^M\right)(\mathbf{x})\right]\right\}\right) \\
&\qquad = \prod_{a,\mathbf{x}}\left(\frac{\delta^3x}{2\pi N_A}\sqrt{\frac{2\pi}{i\delta t\delta^3x}}\right)^3 \exp\left\{\frac{i\delta t\delta^3x}{2}\left[\dot{\mathbf{A}}_a^M(\mathbf{x}) + \left(\overline{\mathbf{D}}^M\overline{\varphi}^M\right)_a(\mathbf{x})\right]^2\right\} \\
&\qquad\qquad = \exp\left[i\delta t\int d^3x\frac{1}{2}\left(\mathbf{E}^M\right)^2\right]\prod_{a,\mathbf{x}}\left(\frac{\delta^3x}{2\pi i\delta t N_A^2}\right)^{3/2} \qquad \text{(III.5.1)}
\end{aligned}$$

and rewrite the class operator as

$$\langle\Phi|C_\alpha|\Psi\rangle = \mathcal{N}\int \mathcal{D}^4 A^{J+1}\Phi^*\big[\mathbf{A}^{J+1};t''\big)$$
$$\times\left(\prod_{M=J}^{0}\int \mathcal{D}^4 A^M \exp\left\{i\delta t\int d^3x\frac{1}{2}\left[\left(\mathbf{E}^M\right)^2-\left(\mathbf{B}^M\right)^2\right]\right\}\right)$$
$$\times\left(\prod_{M=0}^{J+2}\delta\left[G^M\right]\Delta_{G^M}\right)\Psi\left[\mathbf{A}^0;t'\right)e_\alpha[A], \tag{III.5.2}$$

where $\mathbf{E}^M$ and $\mathbf{B}^M$ are as given in (III.3.47) and we have defined the normalization constant

$$\mathcal{N}=\left[\prod_{a,\mathbf{x}}\left(\frac{\delta^3x}{2\pi i\delta t N_A^2}\right)^{3/2}\right]^{J+1}. \tag{III.5.2a}$$

Looking at (III.5.2), we see that it is a lattice realization of the formal expression (III.3.5), with the measure for the configuration space path integral explicitly calculated.

III.5.3 Coarse graining by values of the constraints

Factoring the class operators

In configuration space, the constraint becomes

$$Q=-\mathbf{D}\cdot\mathbf{E}=\mathbf{D}\cdot\dot{\mathbf{A}}+\mathbf{D}^2\varphi=0 \tag{III.5.3}$$

and the electric field part of the Lagrangian is

$$\int d^3x\frac{1}{2}\mathbf{E}^2=\int d^3x\frac{1}{2}\left[\dot{\mathbf{A}}^2+2\dot{\mathbf{A}}\cdot\mathbf{D}\varphi+(\mathbf{D}\varphi)^2\right]. \tag{III.5.4}$$

This is most fruitfully simplified by a new gauge, which we dub the “dotted Coulomb gauge” (DCG),[12] in which

$$\mathbf{D}\cdot\dot{\mathbf{A}}=0. \tag{III.5.5}$$

This differs from the Coulomb gauge in which $\mathbf{D}\cdot\mathbf{A}=0$ because the time derivative does not commute with the covariant gradient $\mathbf{D}$. In this gauge, the constraint becomes $Q=\mathbf{D}^2\varphi=0$ and, after integrating by parts, the electric field part of the Lagrangian becomes

$$\int d^3x\frac{1}{2}\left[\dot{\mathbf{A}}^2+(\mathbf{D}\varphi)^2\right]. \tag{III.5.6}$$

The lattice realization of the gauge condition is

$$G^M=\overline{\mathbf{D}}^M\cdot\dot{\mathbf{A}}^M=0 \tag{III.5.7}$$

[12] It is possible to show that we can always reach this gauge, via an argument analogous to that used in [42] to show that one can always reach the Coulomb gauge in a NAGT.

for $M = 0$ to J, which, taken along with $G^{J+2} = \overline{\varphi}^{J+1} = 0$, leaves one hypersurface worth of gauge conditions G^{J+1} to be specified. We assume that this is defined on some hypersurface away from the region examined by our coarse graining, and ignore it. (It is conventional to assume that it has been used to ensure that the scalar potential vanishes at spatial infinity.) The Fadeev-Popov determinant for the DCG can be calculated from (III.3.2b) to give

$$\Delta_G = \det[-\mathbf{D}^2\partial_t] = \det[-\mathbf{D}^2]\det[\partial_t] \tag{III.5.8}$$

and the class operator is now

$$\begin{aligned}
\langle\Phi|C_\alpha|\Psi\rangle = \mathcal{N}\int \mathcal{D}^3A^{J+1}\det[\partial_t]\Phi^*\left[\mathbf{A}^{J+1};t''\right)\Psi\left[\mathbf{A}^0;t'\right)& \\
\times\left(\prod_{M=J}^{0}\int\mathcal{D}^3A^M\exp\left\{i\delta t\int d^3x\frac{1}{2}\left[\left(\dot{\mathbf{A}}^M\right)^2-\left(\mathbf{B}^M\right)^2\right]\right\}\delta\left[\overline{\mathbf{D}}^M\cdot\dot{\mathbf{A}}^M\right]\det\left[-\left(\overline{\mathbf{D}}^M\right)^2\partial_t\right]\right)& \\
\times\left\{\prod_{M=J}^{0}\int\mathcal{D}\overline{\varphi}^M\exp\left[i\delta t\int d^3x\frac{1}{2}\left(\overline{\mathbf{D}}^M\overline{\varphi}^M\right)^2\right]\right\}\delta\left[G^{J+1}\right]\Delta_{G^{J+1}}e_\alpha[\mathbf{A},\varphi].& \quad \text{(III.5.9)}
\end{aligned}$$

The expression (III.5.9) is beginning to factor into two pieces: a piece depending on the initial and final wave functionals which involves only the vector potential, and a piece describing the coarse graining which involves only the scalar potential. The two factors which still involve both potentials are the $(\mathbf{D}\varphi)^2$ term in the exponential and the indicator functional $e_\alpha[\mathbf{A},\varphi]$. We would like to solve the first problem by changing variables from φ to $\mathbf{D}\varphi$ in the path integral, but the latter is a vector while the former is only a scalar. To construct a scalar corresponding to $\mathbf{D}\varphi$, we need to develop some notation.

First, for the remainder of this section, it will be useful to consider all unadorned variables to be scalars rather than four-vectors. For example, $k = |\mathbf{k}| = \sqrt{\mathbf{k}\cdot\mathbf{k}}$. Now, we define a nonlocal scalar operator $\nabla = (\boldsymbol{\nabla}^2)^{1/2}$ via a Fourier transform:

$$\nabla f(\mathbf{x}) = \int\frac{d^3x'\,d^3k}{(2\pi)^3}e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}ikf(\mathbf{x}') \tag{III.5.10}$$

so that $\nabla^2 = \boldsymbol{\nabla}^2$ is the Laplacian. (We could have defined the square root to have the opposite sign, but it would not substantially change what follows.) Building on the properties of this operator, we define an analogous square root for the covariant Laplacian

$$\mathbf{D}^2 = \boldsymbol{\nabla}^2 + ig(\mathbf{A}^a\cdot\boldsymbol{\nabla}+\boldsymbol{\nabla}\cdot\mathbf{A}^a)T_a - g^2(\mathbf{A}^a\cdot\mathbf{A}^b)T_aT_b \tag{III.5.11}$$

via an expansion (the convergence of which we do not address) as follows:

$$\begin{aligned}
D = (\mathbf{D}^2)^{1/2} &= \left[\boldsymbol{\nabla}^2+(\mathbf{D}^2-\boldsymbol{\nabla}^2)\right]^{1/2} = (\boldsymbol{\nabla}^2)^{1/2}\left[1+\nabla^{-2}(\mathbf{D}^2-\boldsymbol{\nabla}^2)\right]^{1/2} \\
&= \nabla\sum_{n=0}^{\infty}b_n\left[\nabla^{-2}(\mathbf{D}^2-\boldsymbol{\nabla}^2)\right]^n,
\end{aligned} \tag{III.5.12}$$

where $\{b_n\}$ are the Taylor expansion coefficients of $(1+x)^{1/2}$ and ∇^{-2} is another non-local operator

$$\nabla^{-2} f(\mathbf{x}) = \int \frac{d^3x' d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} \left(-\frac{1}{\mathbf{k}^2}\right) f(\mathbf{x}') \tag{III.5.13}$$

defined so that $\boldsymbol{\nabla}^2 \nabla^{-2} = \nabla^{-2}\boldsymbol{\nabla}^2 = 1$.

Now we want to massage the scalar potential part of the action so that it involves $D\varphi$ rather than $\mathbf{D}\varphi$. Integrating by parts, we see that

$$\frac{1}{2}\int d^3x (\mathbf{D}\varphi)^2 = -\frac{1}{2}\int d^3x\, \varphi\, \mathbf{D}\cdot\mathbf{D}\,\varphi = -\frac{1}{2}\int d^3x\, \varphi D^2 \varphi \tag{III.5.14}$$

and now we need to move one of the D operators back to the left. It is straightforward to show (by expanding α and β in Fourier transforms) that

$$\int d^3x\, \alpha(\mathbf{x}) \nabla \beta(\mathbf{x}) = \int d^3x (\nabla\alpha)(\mathbf{x})\beta(\mathbf{x}) \tag{III.5.15}$$

(which is the opposite sign from the integration by parts involving $\boldsymbol{\nabla}$) and

$$\int d^3x\, \alpha(\mathbf{x}) \nabla^{-2} \beta(\mathbf{x}) = \int d^3x (\nabla^{-2}\alpha)(\mathbf{x})\beta(\mathbf{x}). \tag{III.5.16}$$

Using those two results, along with

$$\int d^3x\, \alpha(\mathbf{x})(\mathbf{D}^2 - \boldsymbol{\nabla}^2)\beta(\mathbf{x}) = \int d^3x \left[(\mathbf{D}^2 - \boldsymbol{\nabla}^2)\alpha(\mathbf{x})\right]\beta(\mathbf{x}) \tag{III.5.17}$$

(which follows from the integration by parts procedures for $\boldsymbol{\nabla}$ and $\mathbf{D}$) one can show that

$$-\frac{1}{2}\int d^3x\, \varphi D^2 \varphi = -\frac{1}{2}\int d^3x (D\varphi)^2. \tag{III.5.18}$$

If we define[13]

$$\mathcal{E}^M = i\overline{D}^M \overline{\varphi}^M, \tag{III.5.19}$$

the class operator becomes

$$\begin{aligned}\langle\Phi|C_\alpha|\Psi\rangle &= \mathcal{N}\int \mathcal{D}^3 A^{J+1} \det[\partial_t]\Phi^*\left[\mathbf{A}^{J+1}; t''\right)\Psi\left[\mathbf{A}^0; t'\right) \\ &\times \left(\prod_{M=J}^{0}\int \mathcal{D}^3 A^M \exp\left\{i\delta t \int d^3x \frac{1}{2}\left[\left(\dot{\mathbf{A}}^M\right)^2 - \left(\mathbf{B}^M\right)^2\right]\right\}\delta\left[\overline{\mathbf{D}}^M \cdot \dot{\mathbf{A}}^M\right]\det\left[i\overline{D}^M \partial_t\right]\right) \\ &\qquad \times \left\{\prod_{M=J}^{0}\int \mathcal{D}\mathcal{E}^M \exp\left[i\delta t\int d^3x \frac{1}{2}\left(\mathcal{E}^M\right)^2\right]\right\}\delta\left[G^{J+1}\right]\Delta_{G^{J+1}} e_\alpha[\mathbf{A},\mathcal{E}].\end{aligned} \tag{III.5.20}$$

[13] The factor of i is necessary to make $\mathcal{E}(\mathbf{x})$ a real quantity. In E&M, this change of variables is just changing to $\mathcal{E} = -iE_L$ where E_L is the (scalar) longitudinal part of the electric field.

Now the only obstacle to factorization of the class operator is the indicator functional e_α. If we coarse grain by some temporal and spatial average $\langle Q\rangle$ of the constraint

$$Q = \mathbf{D}^2\varphi = -iD\mathcal{E}, \tag{III.5.21}$$

the indicator functional for a class in which this average lies in the range Δ (which, since Q is an isovector, is a region in isospace which is mapped onto itself by gauge transformations) is (letting $n = \sum_a 1$ be the dimension of the adjoint representation of the gauge group, and keeping in mind that f is a complex isovector quantity)

$$e_\Delta[\langle Q\rangle] = \int\limits_\Delta d^{2n}f\,\delta(f - \langle -iD\mathcal{E}\rangle), \tag{III.5.22}$$

which depends on $\mathbf{A}$ via the operator D. If, however, we coarse grain by values of $iD^{-1}Q$, which classically should vanish whenever Q does, we are coarse graining by $\mathcal{E}$, e_Δ is independent of $\mathbf{A}$, and we can perform the following manipulation:

$$\begin{aligned}\langle\Phi|C_\Delta|\Psi\rangle &= \mathcal{N}\mathcal{C}_\Delta \int \mathcal{D}^3A^{J+1}\det[\partial_t]\Phi^*\left[\mathbf{A}^{J+1};t''\right)\Psi\left[\mathbf{A}^0;t'\right) \\ &\times\left(\prod_{M=J}^{0}\int\mathcal{D}^3A^M\exp\left\{i\delta t\int d^3x\frac{1}{2}\left[\left(\dot{\mathbf{A}}^M\right)^2-\left(\mathbf{B}^M\right)^2\right]\right\}\delta\left[\overline{\mathbf{D}}^M\cdot\dot{\mathbf{A}}^M\right]\det\left[i\overline{D}^M\partial_t\right]\right) \\ &\times\delta\left[G^{J+1}\right]\Delta_{G^{J+1}},\end{aligned} \tag{III.5.23}$$

where

$$\mathcal{C}_\Delta = \left\{\prod_{M=J}^{0}\int\mathcal{D}\mathcal{E}^M\exp\left[i\delta t\int d^3x\frac{1}{2}\left(\mathcal{E}^M\right)^2\right]\right\}e_\Delta[\mathcal{E}]. \tag{III.5.24}$$

This means that

$$D(\Delta,\Delta') = \frac{\mathcal{C}^*_\Delta\mathcal{C}_{\Delta'}}{|\mathcal{C}_u|^2} \tag{III.5.25}$$

and we can apply an argument from section VI.4 of [24]: when the decoherence functional factors in this way, the only way the off-diagonal elements can vanish is if only one of the $\{\mathcal{C}_\Delta\}$ is non-zero. In that case, one diagonal element of the decoherence functional is unity and all the others vanish, which corresponds to a definite prediction of that alternative (100% probability). Thus we have the result: *coarse grainings of* $iD^{-1}(-\mathbf{D}\cdot\mathbf{E})$ *in configuration space fall into two categories: either they yield a definite prediction of a single alternative, or they fail to decohere.* In the former case, we expect that the predicted alternative will be the one consistent with constraint $Q = 0$, but this argument itself does not settle the issue. However, the conjecture seems very likely given that the integrand in the expression (III.5.24) for $\mathcal{C}_\Delta$ is stationary about $\mathcal{E} = 0$, which would seem to make the alternative including $\mathcal{E} = 0$ the one most likely to have a non-vanishing $\mathcal{C}_\Delta$.

A decohering example

We now present explicit calculation of $\mathcal{C}_\Delta$ for one choice of the average $\langle\mathcal{E}\rangle$ which verifies both that coarse grainings of the first class exist and that the alternative predicted is (in this case) indeed the one consistent with the constraint. The demonstration is analogous to that used in section VI.4 of [24] for E&M, and the specialization of the present result (III.A.16) to the Abelian case is in fact a more accurate version of equation (VI.4.12) therein.

We coarse grain by an average $\langle\mathcal{E}\rangle$ over modes so that the indicator functional is

$$e_\Delta[\mathcal{E}] = \int\limits_\Delta d^{2n}f\delta(f - \langle\mathcal{E}\rangle). \tag{III.5.26}$$

The average $\langle\mathcal{E}\rangle$ is defined to be over a time interval Δt and a group of modes in spatial frequency space $\Delta^3 k$. We refer to this group of modes as Ω, which we also use for the mode volume ($\Omega = \Delta t\Delta^3 k$), so that the average is

$$\langle\mathcal{E}\rangle = \frac{1}{\Omega}\int\limits_\Omega dt\, d^3k\, \mathcal{E}_{\mathbf{k}}(t) = \frac{1}{\Omega}\sum_{M\in\Omega}\delta t\int\limits_\Omega d^3k\, \mathcal{E}^M_{\mathbf{k}}, \tag{III.5.27}$$

where $\mathcal{E}_{\mathbf{k}}$ is the Fourier transform

$$\mathcal{E}^M_{\mathbf{k}} = \int \frac{d^3x}{(2\pi)^{3/2}} e^{-i\mathbf{k}\cdot\mathbf{x}}\mathcal{E}^M(\mathbf{x}). \tag{III.5.28}$$

For this coarse graining, the calculation in appendix III.A gives

$$\mathcal{C}_\Delta = \mathcal{K}'\int\limits_\Delta d^{2n}fe^{i\Omega|f|^2}, \tag{III.A.16}$$

where $\mathcal{K}'$ is a constant. The integrand is an "imaginary Gaussian" of width $1/\sqrt{2\Omega}$; For $|f|^2 \gtrsim 1/2\Omega$, the integrand will oscillate rapidly and the contributions to the integral will cancel out. This means that if we average over a large enough group of modes Ω that the region

$$\left|\frac{1}{\Omega}\int\limits_\Omega dt\, d^3k\, \mathcal{E}_{\mathbf{k}}(t)\right|^2 \lesssim \frac{1}{2\Omega} \tag{III.5.29}$$

is contained in a single bin Δ, that will correspond to the only non-negligible $\mathcal{C}_\Delta$, and we will have a definite prediction that the configuration space constraint is satisfied to that accuracy.

This result is less comforting than the Abelian one, since our alternatives were defined not by the usual configuration space constraint $Q = -\mathbf{D}\cdot\mathbf{E}$ but a nonlocal function of it. In E&M, no one would object to analogously coarse graining by the longitudinal component of the electric field rather than its divergence, but in that case the relationship between them does not involve the other components of $\mathbf{A}$ so a similar factorization can be performed on $-\boldsymbol{\nabla}\cdot\mathbf{E}$ as on E_L. However, in a NAGT, coarse graining by $Q = -iD\mathcal{E}$ tangles up $\mathcal{E}$ and $\mathbf{A}$. Even in E&M, we run into this problem if we coarse grain by quantities which involve both E_L and $\mathbf{A}_T$. We examine one such coarse graining in the next section.

III.5.4 Coarse graining E&M by quantities proportional to the constraint

We showed in the previous section that coarse grainings by values of $\mathcal{E}$ in a NAGT (or E_L in E&M) could only decohere in cases where they led to a definite prediction. The demonstration does not work for coarse grainings by $-iD\mathcal{E}$ (or $f[E_L, \mathbf{B}]$ in E&M). We will now exhibit such a coarse graining in E&M which decoheres, but predicts non-zero probabilities for more than one alternative, thus verifying that the property described in the previous section does not always hold.

The situation for this coarse graining is reminiscent of the physical decoherence discussed in Chapter IV, in which "system" of interest is coupled to an "environment" which is not measured, but carries away phase information which causes sets of alternatives describing the "system" to decohere, but with the following differences. The "system" variables $\mathbf{A}$ are coupled to the "environment" variables $\mathcal{E}$ not by the action, but by the coarse graining itself, and here it is the initial state rather than the coarse graining which is independent of the "environment" $\mathcal{E}$. But as we shall see, this is still a mechanism which can produce decoherence of a sort different than that seen in the previous section, and lead to more than one alternative having non-zero probability. For our purposes, it will be most useful to consider coarse grainings by functionals of E_L and $\mathbf{B}$ in the Abelian gauge theory of electromagnetism. (Although we will briefly mention, at the end of the section, a similar result in another theory to illustrate the generality of the mechanism described here.)

In the Abelian theory, the dotted Coulomb gauge (III.5.5) is equivalent to the Coulomb gauge $A_L = 0$ and (III.5.9) becomes

$$\begin{aligned}\langle\Phi|C_\alpha|\Psi\rangle = \mathcal{N}\int &\mathcal{D}^2 A_T^{J+1} \det[\partial_t]\Phi^*\left[\mathbf{A}_T^{J+1};t''\right)\Psi\left[\mathbf{A}_T^0;t'\right) \\ &\times\left(\prod_{M=J}^{0}\int\mathcal{D}^2 A_T^M \exp\left\{i\delta t\int d^3x\frac{1}{2}\left[\left(\dot{\mathbf{A}}_T^M\right)^2-\left(\boldsymbol{\nabla}\times\dot{\mathbf{A}}_T^M\right)^2\right]\right\}\det[-\nabla]\right) \\ &\times\left\{\prod_{M=J}^{0}\int\mathcal{D}\overline{\varphi}^M \exp\left[i\delta t\int d^3x\frac{1}{2}\left(\boldsymbol{\nabla}\overline{\varphi}^M\right)^2\right]\right\} e_\alpha[\mathbf{A}_T,\varphi].\end{aligned} \tag{III.5.30}$$

For the purposes of E&M, we need not concern ourselves with the details of the lattice approximation, because the reduced Hamiltonian is free from the operator ordering ambiguities discussed in section III.3.4.[14] This means that any choice of operator ordering convention gives the same reduced Hamiltonian, and in light of the discussion in footnote 11, page 70 this means that different lattice realizations of the path integral will be equivalent. We are

[14] For example, in the axial gauge (III.3.13), the functional defined in (III.3.19) becomes $\pi_n[\mathbf{A}_\perp, \boldsymbol{\pi}_\perp] = -\partial_n^{-1}\boldsymbol{\nabla}_\perp\cdot\boldsymbol{\pi}_\perp$, which is independent of $\mathbf{A}_\perp$ so that $\frac{1}{2}(\pi_n[\mathbf{A}_\perp, \boldsymbol{\pi}_\perp])^2$ can be unambiguously converted into an operator expression.

thus justified in working with the formal equivalent of (III.5.30):

$$\langle\Phi|C_\alpha|\Psi\rangle = \Delta_G \int\limits_\alpha \mathcal{D}^2 A_T\, \mathcal{D}\varphi\, \Phi^*\left[\mathbf{A}_T^{J+1};t''\right) \times \exp\left\{i\int d^4x \frac{1}{2}\left[\left(\dot{\mathbf{A}}_T\right)^2 - \left(\boldsymbol{\nabla}\times\dot{\mathbf{A}}_T\right)^2 + (\boldsymbol{\nabla}\varphi)^2\right]\right\}\Psi\left[\mathbf{A}_T^0;t'\right). \quad \text{(III.5.31)}$$

Since we can observe that the "physical degrees of freedom" upon which the wave functionals depend are just the transverse components of $\mathbf{A}$, it is useful to factor out the wave functionals and write

$$\langle\Phi|C_\alpha|\Psi\rangle = \int \mathcal{D}^2 A''_T\, \mathcal{D}^2 A'_T\, \Phi^*[\mathbf{A}''_T;t'')\, C_\alpha[\mathbf{A}''_T;t''|\mathbf{A}'_T;t')\Psi\left[\mathbf{A}'_T;t'\right), \quad \text{(III.5.32)}$$

where

$$C_\alpha[\mathbf{A}''_T;t''|\mathbf{A}'_T;t') = \Delta_G \int\limits_{\mathbf{A}''_T\alpha\mathbf{A}'_T} \mathcal{D}^2 A_T\, \mathcal{D}\varphi\, e^{i\int_{t'}^{t''}dtL}. \quad \text{(III.5.33)}$$

If we write the initial and final conditions as

$$\rho''_{t''}[\mathbf{A}''_{T2},\mathbf{A}''_{T1}] = \sum_i \Phi_i^*[\mathbf{A}''_{T2};t'')\, p''_i \Phi_i\left[\mathbf{A}''_{T1};t''\right) \quad \text{(III.5.34a)}$$

$$\rho'_{t'}[\mathbf{A}'_{T2},\mathbf{A}'_{T1}] = \sum_j \Psi_j^*[\mathbf{A}'_{T2};t')\, p'_j \Psi_j\left[\mathbf{A}'_{T1};t'\right), \quad \text{(III.5.34b)}$$

we have, from (III.5.32) and (I.3.12),

$$D(\alpha,\alpha') \propto \int \mathcal{D}^2 A''_{T2}\mathcal{D}^2 A''_{T1}\mathcal{D}^2 A'_{T2}\mathcal{D}^2 A'_{T1} \times \rho''_{t''}[\mathbf{A}''_{T2},\mathbf{A}''_{T1}]C_\alpha[\mathbf{A}''_{T1};t''|\mathbf{A}'_{T1};t')\rho'_{t'}[\mathbf{A}'_{T1},\mathbf{A}'_{T2}]C^*_{\alpha'}[\mathbf{A}''_{T2};t''|\mathbf{A}'_{T2};t'), \quad \text{(III.5.35)}$$

where we have established the useful convention that $\propto$ indicates a proportionality constant which is the same for all classes and thus can be absorbed into the normalization. The quantity we choose to define our alternatives is

$$g[\mathbf{B}^\iota]\left|\langle E_L\rangle\right|^2, \quad \text{(III.5.36)}$$

where $g[\mathbf{B}^\iota]$ is a functional (which we take to be positive semidefinite for reasons to become clear later) of the magnetic field configuration $\mathbf{B}^\iota$ on some time slice t_i, and $\langle\cdot\rangle$ indicates an average over some mode volume (i.e., an average over wavenumber and time[15]). The indicator function for this quantity to lie in some interval Δ is

$$e_\Delta = \int\limits_\Delta df\,\delta\Big(f - g[\mathbf{B}^\iota]\left|\langle E_L\rangle\right|^2\Big) = \int\limits_\Delta df\int db\,da\,\delta(f-ba)\delta(b-g[\mathbf{B}^\iota])\delta\Big(a - \left|\langle E_L\rangle\right|^2\Big), \quad \text{(III.5.37)}$$

[15]Since it involves a time average, this sort of alternative is not accessible in a standard operator-and-state formulation of quantum mechanics.

which allows us to write

$$
\begin{aligned}
C_\Delta[\mathbf{A}''_T;t''|\mathbf{A}'_T;t') &= \int\limits_\Delta df \int db\, da \delta(f-ba)\mathcal{A}(a)\mathcal{B}\left[\mathbf{A}''_T,\mathbf{A}'_T,b\right) \\
&= \int\limits_\Delta df \int \frac{db}{|b|}\mathcal{A}\left(\frac{f}{b}\right)\mathcal{B}\left[\mathbf{A}''_T,\mathbf{A}'_T,b\right),
\end{aligned}
\tag{III.5.38}
$$

where

$$
\mathcal{A}(a) = \int \mathcal{D}\varphi \exp\left[i\int_{t'}^{t''} dt \int d^3x \frac{1}{2}\left(\boldsymbol{\nabla}\varphi\right)^2\right]\delta\left(a - |\langle\nabla\varphi\rangle|^2\right)\det[-\nabla] \tag{III.5.38a}
$$

and

$$
\mathcal{B}\left[\mathbf{A}''_T,\mathbf{A}'_T;b\right) = \int\limits_{\mathbf{A}''_T\mathbf{A}'_T} \mathcal{D}^2A_T \exp\left\{i\int_{t'}^{t''} dt \int d^3x \frac{1}{2}\left[\left(\dot{\mathbf{A}}_T\right)^2 - \left(\boldsymbol{\nabla}\times\mathbf{A}_T\right)^2\right]\right\}\delta(b - g[\boldsymbol{\nabla}\times\mathbf{A}^\iota_T]). \tag{III.5.38b}
$$

Writing the average over a group of modes as

$$
\langle\nabla\varphi\rangle = \frac{1}{\Omega}\int\limits_\Omega dt\, d^3k[ik\overline{\varphi}_{\mathbf{k}}(t)], \tag{III.5.39}
$$

a calculation analogous to the one in appendix III.A tells us that

$$
\begin{aligned}
\mathcal{A}(a) &\propto \int \mathcal{D}\Upsilon \exp\left(i\Omega\sum_{\sigma\in\Omega}|\Upsilon_\sigma|^2\right)\delta\left(a - |\Upsilon_0|^2\right) \\
&\propto \int d\Upsilon_0^{\mathrm{R}} d\Upsilon_0^{\mathrm{I}} \exp\left(i\Omega\,|\Upsilon_0|^2\right)\delta\left(a - |\Upsilon_0|^2\right) \propto e^{i\Omega a}\Theta(a),
\end{aligned}
\tag{III.5.40}
$$

where $\Theta(a)$ is the Heaviside step function. Meanwhile, we can write $\mathcal{B}$ as[16]

$$
\mathcal{B}\left[\mathbf{A}''_T,\mathbf{A}'_T;b\right) = \int \mathcal{D}^2A^\iota_T\, \mathcal{G}\left[\mathbf{A}''_T;t''|\mathbf{A}^\iota_T;t_i\right)\delta(b - g[\boldsymbol{\nabla}\times\mathbf{A}^\iota_T])\,\mathcal{G}\left[\mathbf{A}^\iota_T;t_i|\mathbf{A}'_T;t'\right), \tag{III.5.41}
$$

where

$$
\mathcal{G}\left[\mathbf{A}''_T;t''|\mathbf{A}'_T;t'\right) = \int\limits_{\mathbf{A}''_T\mathbf{A}'_T} \mathcal{D}^2A_T \exp\left\{i\int_{t'}^{t''} dt \int d^3x \frac{1}{2}\left[\left(\dot{\mathbf{A}}_T\right)^2 - \left(\boldsymbol{\nabla}\times\mathbf{A}_T\right)^2\right]\right\} \tag{III.5.42}
$$

[16] The attentive reader may notice that we are implicitly expressing our coarse graining in terms of $\mathbf{A}^\iota$ as though it corresponded to $\mathbf{A}$ on a single lattice slice $\left(\mathbf{A}^I\right)$ rather than an average $\left(\overline{\mathbf{A}}^I\right)$, as we were instructed to do in section III.3.3. This is not a problem, because, as discussed earlier, the operator ordering ambiguities that led us to make the distinction between the two are not present in E&M.

is the propagator for the $\mathbf{A}_T$ sector of the theory. By Fourier transforming the spatial dependence of $\mathbf{A}_T$, $\mathcal{G}$ can be seen to be the propagator for a harmonic oscillator whose natural frequency depends on the wave number $\mathbf{k}$ of the mode. Equation (III.5.41) allows us to write the dependence implied by (III.5.35) and (III.5.38) of $D(\Delta_2, \Delta_1)$ on the initial and final conditions as

$$\int \mathcal{D}^2 A''_{T2} \mathcal{D}^2 A''_{T1} \mathcal{D}^2 A'_{T2} \mathcal{D}^2 A'_{T1} \rho''_{t''}[\mathbf{A}''_{T2}, \mathbf{A}''_{T1}] \mathcal{B}\,[\mathbf{A}''_{T1}, \mathbf{A}'_{T1}; b_1)\, \rho'_{t'}[\mathbf{A}'_{T1}, \mathbf{A}'_{T2}] \mathcal{B}^*[\mathbf{A}''_{T2}, \mathbf{A}'_{T2}; b_2)$$
$$= \int \mathcal{D}^2 A^\iota_{T2} \mathcal{D}^2 A^\iota_{T1} \rho''_{t_i}[\mathbf{A}^\iota_{T2}, \mathbf{A}^\iota_{T1}] \delta(b_1 - g[\boldsymbol{\nabla} \times \mathbf{A}^\iota_{T1}]) \rho'_{t_i}[\mathbf{A}^\iota_{T1}, \mathbf{A}^\iota_{T2}] \delta(b_2 - g[\boldsymbol{\nabla} \times \mathbf{A}^\iota_{T2}]), \tag{III.5.43}$$

where we have used the propagator $\mathcal{G}$ to transform $\rho'_{t'}$ to ρ'_{t_i} and $\rho''_{t''}$ to ρ''_{t_i}. If the final state is one of future indifference:

$$\rho''[\mathbf{A}''_{T2}, \mathbf{A}''_{T1}] \propto \delta[\mathbf{A}''_{T2} - \mathbf{A}''_{T1}] \tag{III.5.44}$$

(which is preserved by the propagator), (III.5.43) becomes proportional to

$$\int \mathcal{D}^2 A^\iota_T\, \delta(b_1 - g[\boldsymbol{\nabla} \times \mathbf{A}^\iota_T]) \rho'_{t_i}[\mathbf{A}^\iota_T, \mathbf{A}^\iota_T] \delta(b_2 - g[\boldsymbol{\nabla} \times \mathbf{A}^\iota_T]) = \delta(b_2 - b_1)\, p(b_1), \tag{III.5.45}$$

where

$$p(b) = \int \mathcal{D}^2 A^\iota_T\, \delta(b - g[\boldsymbol{\nabla} \times \mathbf{A}^\iota_T]) \rho'_{t_i}[\mathbf{A}^\iota_T, \mathbf{A}^\iota_T]. \tag{III.5.46}$$

Combining (III.5.35) and (III.5.38) with the expression for $\mathcal{A}$ in (III.5.40) and this result concerning $\mathcal{B}$, we have

$$D(\Delta, \Delta') \propto \int\limits_{\Delta} df \int\limits_{\Delta'} df' \int \frac{db}{b^2} p(b) e^{i\Omega(f-f')/b}\, \Theta\left(\frac{f}{b}\right) \Theta\left(\frac{f'}{b}\right). \tag{III.5.47}$$

With the condition that $g[\mathbf{B}^\iota]$ is everywhere non-negative, we see from (III.5.46) that $p(b)$ vanishes for negative b and the step functions above become $\Theta(f)\Theta(f')$. If we define the regions $\{\Delta\}$ to cover the positive real axis, we can drop the step functions to give

$$D(\Delta, \Delta') \propto \int\limits_{\Delta} df \int\limits_{\Delta'} df'\, G(f - f'), \tag{III.5.48}$$

where

$$G(y) = \int_0^\infty \frac{db}{b^2} p(b) e^{i\Omega y/b}. \tag{III.5.48a}$$

Note that since $G(f - f')$ depends only on the difference between f and f', no value of f is preferred over any other. In particular, if the bins $\{\Delta\}$ are all the same size, $D(\Delta, \Delta')$ depends only on the relative separation of Δ and Δ', not their absolute location. This means that if there is decoherence, (III.5.48) predicts that the measured quantity is equally likely to have any value.

It is possible to choose the $p(b)$ (which is determined by the initial conditions) to produce at least weak decoherence. For example, let $p(b)$ be a Gaussian in $1/b$:

$$p(b) = A_\lambda \Theta(b)\Theta(\lambda^{-1} - b)e^{-1/2\sigma^2 b^2}, \tag{III.5.49}$$

where A_λ is a cutoff-dependent normalization given by

$$A_\lambda^{-1} = \int_0^{\lambda^{-1}} db\, e^{-1/2\sigma^2 b^2} < \frac{\sigma\sqrt{2\pi}}{\lambda^2} \tag{III.5.49a}$$

to ensure $\int_0^\infty p(b) = 1$.

If λ is small enough, the leading terms in the decoherence functional will not depend on it. If the real parts of the off-diagonal elements of the decoherence functional are much less than the diagonal elements, the coarse graining will exhibit approximate weak decoherence [*cf.* (I.3.5)] The calculation in appendix III.C shows for bins of equal size Δ that, to lowest order in $e^{-(\Omega\sigma\Delta)^2/2}$,

$$\frac{\operatorname{Re} D(J+\Delta J, J)}{D(J,J)} \lesssim \frac{\exp[-(\Omega\sigma\Delta)^2(|\Delta J|-1)^2/2]}{\Omega\sigma\Delta\sqrt{2\pi}}. \tag{III.5.50}$$

So $D(J, J\pm 1)$ is suppressed by a factor of $(\Omega\sigma\Delta)^{-1}$ relative to $D(J,J)$, while all the other elements of the decoherence functional are exponentially suppressed. In general, we expect this sort of result if $\operatorname{Re} G(y)$ falls off on a scale which is small compared to Δ [which should in general be determinable from a steepest descents evaluation of (III.5.48a).] Schematically (Fig. III.1), if $\operatorname{Re} G(y)$ becomes negligible for $|y| \gtrsim \delta$, with $\delta \leq \Delta$, the integral for $\operatorname{Re} D(J, J+\Delta J)$ for $|\Delta J| \geq 2$ will include none of the region for which $G(y)$ is significant. The area of the region in the integral for $\operatorname{Re} D(J, J\pm 1)$ for which $G(y)$ is significant is $\delta^2/2$, while that for $D(J,J)$ is $2\Delta\delta - \delta^2$, so $\operatorname{Re} D(J, J\pm 1)$ is suppressed by a factor of δ/Δ.

This means that if $\Omega\sigma\Delta \gg 1$, *this coarse graining by* $g[\mathbf{B}^\iota]\big|\langle E_L\rangle\big|^2$ *decoheres weakly for the initial condition* (III.5.49) *and the final condition of future indifference, and there is an equal probability for the value to fall into any of the evenly spaced bins.* A curious corollary is that if we coarse grain by combining bins 0 through $J_0 - 1$ into one alternative $c_<$, corresponding to $g[\mathbf{B}^\iota]\,|\langle E_L\rangle|^2 < J_0\Delta$, and all the bins J_0 and up into another alternative $c_>$, corresponding to $g[\mathbf{B}^\iota]\big|\langle E_L\rangle\big|^2 > J_0\Delta$, we find that since $p_<$ is a sum of J_0 equal terms and $p_>$ is an infinite sum of the same terms, $p_< = 0$ and $p_> = 1$ for any finite J_0, a definite prediction that $g[\mathbf{B}^\iota]\,|\langle E_L\rangle|^2 > J_0\Delta$. This sort of phenomenon is common in the use of path integral methods (for another example, see [43]) and is related to the non-differentiability of Brownian paths.

Finally, let us comment on the significance of this result. If we coarse grained by values of the corresponding phase space quantity, $g[\mathbf{B}^\iota]\,|\langle\pi_L\rangle|^2$, (III.4.2) would ensure that we found a definite prediction that it vanished. Thus the phase space and configuration space theories make different predictions. There are well-documented examples in generalized quantum mechanics where analogous phase space and configuration space coarse grainings lead to different decoherence functionals (for example, in Sections V.4.2 and VI.4 of [24]), but in those cases, there was a coarse graining by momentum which decohered while coarse

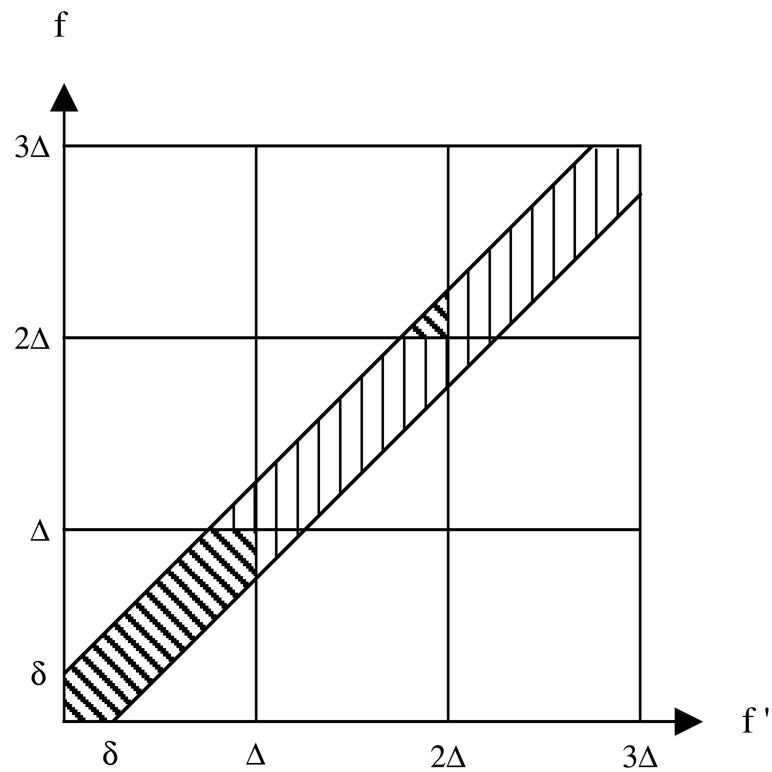


Figure III.1: The regions of integration of a general $G(f - f')$ to produce the decoherence functional $D(J, J')$ via (III.5.48). If $G(y)$ is negligible for $|y| \gtrsim \delta$, integrals of $G(y)$ over regions two or more spots off the diagonal ($|J - J'| \geq 2$) will be negligible. Squares on the diagonal ($J = J'$) have a region of area $2\Delta\delta - \delta^2$ over which $G(y)$ is appreciable. Squares one spot off the diagonal ($|J - J'| = 1$) include some non-negligible values of $G(y)$, but only in a triangular region of area $\delta^2/2$. Thus $D(J, J \pm 1)$ should be suppressed by a factor of δ/Δ relative to $D(J, J)$. Compare Fig. 1 of [12].

graining by the equivalent quantity in terms of velocity did not. When the coarse graining by velocity decohered, it agreed with the coarse graining by momentum. The present result is the first case known to the author of corresponding phase space and configuration space coarse grainings, both of which decohere, but which give conflicting probabilities.

This result is not limited to constrained theories. Another system in which similar phenomena can occur is the non-relativistic quantum mechanics of a free particle with two degrees of freedom and an independent harmonic oscillator. In that case, one coarse grains by the product of some function of the position of the harmonic oscillator at one instant of time with the square of a time average of the velocity of the free particle. If the initial state is a zero-momentum eigenstate of the free particle tensored with a suitable state of the harmonic oscillator, one finds equal probability of any alternative, even though the corresponding phase space coarse graining yields a definite prediction that the quantity vanishes.

These disagreements between configuration space and phase space predictions are another example of classically equivalent theories which yield different predictions upon quantization. Here, the classical equivalence which does not hold is that between momentum and velocity. Possible motivations for choosing between these quantization schemes in the case of gauge theories are discussed in the conclusion (section III.7).

III.6 A few words about Lorentz invariance

Since our implementation of the sum over histories for the generalized quantum mechanics of a NAGT has relied rather heavily on a division into time and space, it is worth mentioning how little the formal theory really does to single out a preferred Lorentz frame. The phase space theory is of course not Lorentz-invariant, as the conjugate momenta are defined with respect to a particular time. Since $-\boldsymbol{\pi}$ and $\mathbf{B}$ are treated differently, it is not possible to combine them into a field strength tensor which Lorentz transforms appropriately. This is the source of the apparent asymmetry between different components of the equations of motion

$$D_\mu G^{\mu\nu} = 0; \tag{III.6.1}$$

the constraints hold identically, while the others do not.

However, the formal configuration space theory (and not just the "physical" configuration space coarse grainings defined in section III.5.1) *can* be cast into a form which is manifestly Lorentz-invariant. In the formal configuration space expression

$$\langle\Phi|C_\alpha|\Psi\rangle = \int\limits_\alpha \mathcal{D}^4A\,\Phi^*[\mathbf{A}'';t'')\delta[G]\Delta_G[A]\exp\left(-i\int_{t'}^{t''}dt\int d^3x\frac{1}{4}G^a_{\mu\nu}G_a^{\mu\nu}\right)\Psi[\mathbf{A}';t'), \tag{III.6.2}$$

$\mathbf{E}$ and $\mathbf{B}$ are treated on equal footing from a spacetime point of view as part of the tensor $G^{\mu\nu}$. Lorentz invariance is broken in two ways, both concerning the initial and final wave functionals Ψ and Φ. First, they are attached on surfaces of constant coördinate time rather than arbitrary spacelike surfaces; second, the operator constraints (III.2.18) on Ψ and Φ treat $\varphi = A^0$ and $\mathbf{A} = A^i\mathbf{e}_i$ unequally. In this section, we demonstrate that these two problems are related to one another, and show how the conditions satisfied by the wave functionals can be related to the surfaces on which they are evaluated.

We can generalize (III.6.2) to arbitrary (spacelike) initial and final surfaces in the straightforward manner:

$$\langle\Phi|C_\alpha|\Psi\rangle = \int\limits_\alpha \mathcal{D}^4A\,\Phi^*\left[A^{(\sigma'')};\sigma''\right)\delta[G]\Delta_G[A]e^{i\int_{\sigma'}^{\sigma''}d^4x\mathcal{L}(x)}\Psi\left[A^{(\sigma')};\sigma'\right), \tag{III.6.3}$$

where $A^{(\sigma)}$ is the restriction of the function $A(x)$ (here the four-vector potential, but the definition will apply to any function defined over spacetime) onto the three-surface σ, and the integral for the action is over the region bounded by σ' and σ''. Using the sum over all histories to define a propagator

$$\mathcal{G}\left[A^{(\sigma'')};\sigma''\middle|\,A^{(\sigma')};\sigma'\right) = \int\limits_{A^{(\sigma'')}A^{(\sigma')}} \mathcal{D}^4A\,\delta[G]\Delta_G[A]e^{i\int_{\sigma'}^{\sigma''}d^4x\mathcal{L}(x)}, \tag{III.6.4}$$

we can go from a wave functional Ψ defined on one spacelike surface to one defined on another:

$$\Psi\left[A^{(\sigma'')};\sigma''\right) = \int \mathcal{D}^4A^{(\sigma')}\mathcal{G}\left[A^{(\sigma'')}\sigma''\middle|\,A^{(\sigma')}\sigma'\right)\Psi\left[A^{(\sigma')};\sigma'\right). \tag{III.6.5}$$

The class operators defined from (III.6.3) for different choices of initial and final surfaces are the same so long as all the spacetime points $\{x\}$ at which the coarse grainings restrict the fields $A(x)$ still lie in between the initial and final surfaces.

The conditions satisfied by the wave functional are a consequence of the gauge invariance of the path integral for the propagator, as discussed in [44].[17] Defining coördinates $\{\xi^i\}$ on the 3-surface σ and specifying its embedding in the flat Minkowski space as $\boldsymbol{\{}x^\mu\{\xi^i\}\boldsymbol{\}}$, the metric induced on the surface will be

$$ds^2 = \eta_{\mu\nu}\frac{\partial x^\mu}{\partial\xi^i}d\xi^i\frac{\partial x^\nu}{\partial\xi^j}d\xi^j = h_{ij}d\xi^i d\xi^j, \tag{III.6.6}$$

$$h_{ij} = \eta_{\mu\nu}\frac{\partial x^\mu}{\partial\xi^i}\frac{\partial x^\nu}{\partial\xi^j}. \tag{III.6.6a}$$

The condition that σ be spacelike means that the three-metric $\{h_{ij}\}$ is positive definite, so that the volume element on σ is

$$d^3\Sigma = d^3\xi\sqrt{h}, \tag{III.6.7}$$

where $h = \det\{h_{ij}\}$. Now the restriction of $A(x)$ onto σ is defined by $A^{(\sigma)}(\xi) = A\boldsymbol{(}x(\xi)\boldsymbol{)}$, and is a function of the three coördinates $\{\xi^i\}$ alone. This is the first argument of the

[17]It should be stressed that we are here considering the conditions imposed by *gauge* invariance on the variation of Ψ with respect to the gauge fields. As discussed in [45, 46], there are also conditions imposed on Ψ (viewed as a function of $A^{(\sigma)}$ and σ) by *Lorentz* invariance. They involve changes in the surface σ. Roughly speaking, variations of σ lying in σ itself correspond to relabelling of the points in the surface, and the function $\Psi(\sigma)$ must change to reflect this. The conditions on variations of σ normal to σ is the Schrödinger equation associated with the the propagation rule (III.6.5).

wave functional $\Psi[A^{(\sigma)};\sigma)$. Equation (III.6.5) shows that the dependence of Ψ on its first argument is the same as the dependence of the propagator $\mathcal{G}$ on *its* first argument. Since the path integral in (III.6.4) is invariant under gauge transformations on A, the propagator must be invariant under the effects of those gauge transformations on $A^{(\sigma')}$ and $A^{(\sigma'')}$. Since the gauge transformation

$$\delta A^a_\mu = -\nabla_\mu \, \delta\Lambda_a - g f^c_{ab} A^c_\mu \delta\Lambda^b \qquad \text{(III.2.1)}$$

is nonlocal, the change in $A^{(\sigma)}$ cannot be described by using only the restriction $\delta\Lambda^{(\sigma)}(\xi) = \delta\Lambda\big(x(\xi)\big)$ of the gauge transformation parameter $\delta\Lambda$ onto the surface σ. To identify the troublesome component of the gradient which introduces values of $\delta\Lambda$ off of σ, it is useful to define a projection tensor

$$\Sigma^\nu_\mu = \frac{\partial\xi^i}{\partial x^\mu}\frac{\partial x^\nu}{\partial\xi^i}, \qquad \text{(III.6.8)}$$

where $\partial\xi^i/\partial x^\mu$ is the gradient of ξ with respect to x lying in σ so that

$$\frac{\partial\xi^i}{\partial x^\mu}\frac{\partial x^\mu}{\partial\xi^j} = \delta^i_j. \qquad \text{(III.6.9)}$$

This also follows from the chain rule

$$\frac{\partial}{\partial\xi^i} = \frac{\partial x^\mu}{\partial\xi^i}\frac{\partial}{\partial x^\mu}. \qquad \text{(III.6.10)}$$

Defining a complementary projection tensor $\Upsilon^\nu_\mu = \delta^\nu_\mu - \Sigma^\nu_\mu$, we wish to project out the components of $\{A^\mu\}$ with Σ and Υ. Since $\operatorname{Tr}\Sigma = 3$ and $\operatorname{Tr}\Upsilon = 1$, it is convenient to define projected objects with the number of components equal to the rank of the corresponding projection. Thus, projections along Σ are more concisely defined by simply projecting with $\partial x^\nu/\partial\xi^i$. Defining[18]

$$\aleph_i(\xi) = \frac{\partial x^\nu}{\partial\xi^i} A^{(\sigma)}_\nu(\xi) \qquad \text{(III.6.11)}$$

and using (III.6.10), the component of (III.2.1) lying in σ is

$$\delta\aleph^a_i = -\frac{\partial\delta\Lambda^{(\sigma)}_a}{\partial\xi^i} - g f^c_{ab}\aleph^c_i\delta\Lambda^{(\sigma)}_b, \qquad \text{(III.6.12)}$$

which is expressed entirely in terms of functions of ξ.

To look at the projection of a four-vector by the rank-one Υ, it is convenient to convert it into a scalar by dotting it into some arbitrary timelike vector v. Hence the component of A out of the surface σ is

$$\phi(\xi) = -v^\mu\Upsilon^\nu_\mu(\xi)A_\nu\big(x(\xi)\big) = -u^\nu(\xi)A_\nu\big(x(\xi)\big). \qquad \text{(III.6.13)}$$

[18] We call this $\aleph_i$ rather than $A^{(\sigma)}_i$ to emphasize that the components $\aleph_1, \aleph_2, \aleph_3$ are defined with respect to the coördinates ξ^1, ξ^2, ξ^3 lying in the surface σ and are not in general the spatial components $A^{(\sigma)}_1, A^{(\sigma)}_2, A^{(\sigma)}_3$ defined with respect to the Cartesian spatial coördinates x^1, x^2, x^3.

Since Υ has rank one, all possible vectors $u^\nu = v^\mu \Upsilon_\mu^\nu$ determined from different v's will be parallel to one another. Since

$$u^\nu \frac{\partial \xi^i}{\partial x^\nu} = 0, \tag{III.6.14}$$

u must be parallel to the normal to the surface σ. (We could choose it to be the normal itself, but the normalization factor will turn out to be irrelevant in what follows.) Taking the dot product of (III.2.1) with u, and defining

$$\frac{\partial}{\partial u} = u^\nu \frac{\partial}{\partial x^\nu}, \tag{III.6.15}$$

we have

$$\delta\phi^a = \left(\frac{\partial \delta\Lambda_a}{\partial u}\right)^{(\sigma)} - g f_{ab}^c \phi^c \delta\Lambda_b^{(\sigma)}, \tag{III.6.16}$$

which cannot be determined from ϕ and $\delta\Lambda^{(\sigma)}$ alone.

Now, since the variation of $\mathcal{G}\left[A^{(\sigma)}\sigma \middle| A^{(\sigma')}\sigma'\right)$ under a gauge transformation must vanish, this must also be true for Ψ. That variation is given in terms of the functional derivatives by

$$\delta\Psi[A^{(\sigma)}] = \int d^3\xi \sqrt{h} \frac{\mathcal{D}\Psi}{\mathcal{D}A_\mu}\delta A_\mu = \int d^3\xi \sqrt{h}\left(\frac{\mathcal{D}\Psi}{\mathcal{D}A_\mu^{(\sigma)}}\Sigma_\mu^\nu \delta A_\nu^{(\sigma)} + \frac{\mathcal{D}\Psi}{\mathcal{D}A_\mu^{(\sigma)}}\Upsilon_\mu^\nu \delta A_\nu^{(\sigma)}\right). \tag{III.6.17}$$

In general, Σ and Υ will depend on the coördinate ξ, but they will still commute with the gauge transformation δ and the functional differentiation $\mathcal{D}$. Put otherwise, the same amount of information is included in $(\{\aleph_i\}, \phi)$ as in $\{A_\mu^{(\sigma)}\} \equiv \left(\mathbf{A}^{(\sigma)}, \varphi^{(\sigma)}\right)$, so that Ψ may be viewed as a functional of $\aleph$ and ϕ, in which case (III.6.17) becomes

$$\begin{aligned}\delta\Psi[\aleph, \phi] &= \int d^3\xi \sqrt{h}\left(\frac{\mathcal{D}\Psi}{\mathcal{D}\aleph_i}\delta\aleph_i + \frac{\mathcal{D}\Psi}{\mathcal{D}\phi}\delta\phi\right) \\ &= \int d^3\xi \sqrt{h}\left[\delta\Lambda_a^{(\sigma)} g f_{ab}^c \left(\aleph_i^c \frac{\mathcal{D}\Psi}{\mathcal{D}\aleph_i^b} + \phi^c \frac{\mathcal{D}\Psi}{\mathcal{D}\phi^b}\right) - \frac{\partial \delta\Lambda_a^{(\sigma)}}{\partial \xi^i}\frac{\mathcal{D}\Psi}{\mathcal{D}\aleph_i^a} + \left(\frac{\partial\delta\Lambda_a}{\partial u}\right)^{(\sigma)} \frac{\mathcal{D}\Psi}{\mathcal{D}\phi^a}\right]. \end{aligned}\tag{III.6.18}$$

Integrating the second term by parts and discarding the term at spatial infinity gives

$$\begin{aligned}\delta\Psi[\aleph, \phi] = \int d^3\xi \sqrt{h}\Bigg\{&\delta\Lambda_a^{(\sigma)}\left[g f_{ab}^c \aleph_i^c \frac{\mathcal{D}\Psi}{\mathcal{D}\aleph_i^b} + g f_{ab}^c \phi^c \frac{\mathcal{D}\Psi}{\mathcal{D}\phi^b} + \frac{1}{\sqrt{h}}\frac{\partial}{\partial \xi^i}\left(\sqrt{h}\frac{\mathcal{D}\Psi}{\mathcal{D}\aleph_i^a}\right)\right] \\ &+ \left(\frac{\partial \delta\Lambda_a}{\partial u}\right)^{(\sigma)} \frac{\mathcal{D}\Psi}{\mathcal{D}\phi^a}\Bigg\}.\end{aligned}\tag{III.6.19}$$

For this to vanish for arbitrary $\delta\Lambda(x)$, the coefficients of $\delta\Lambda^{(\sigma)}(\xi)$ and $(\partial_u\delta\Lambda)^{(\sigma)}(\xi)$ must vanish separately. This leads to the generalization of (III.2.18):

$$\frac{\mathcal{D}}{\mathcal{D}\phi^a}\Psi[\aleph,\phi;\sigma)=0 \tag{III.6.20a}$$

$$\left(\delta_{ab}\frac{1}{\sqrt{h}}\frac{\partial}{\partial\xi^i}\sqrt{h}+gf^c_{ab}\aleph^c_i\right)\frac{\mathcal{D}}{\mathcal{D}\aleph^b_i}\Psi[\aleph,\phi;\sigma)=0. \tag{III.6.20b}$$

Recognizing the form of the *geometric* "covariant divergence" on a curved manifold, we see that the general conditions are

$$\frac{\mathcal{D}}{\mathcal{D}\phi}\Psi[\aleph,\phi;\sigma)=0 \tag{III.6.21a}$$

$$D_i\frac{\mathcal{D}}{\mathcal{D}\aleph_i}\Psi[\aleph,\phi;\sigma)=0, \tag{III.6.21b}$$

where D_i is the "covariant" gradient in both the gauge and geometric senses of the word:

$$(D_i\zeta^j)_a=\frac{\partial\zeta^j_a}{\partial\xi^i}+\Gamma^j_{ik}\zeta^k_a+gf^c_{ab}\aleph^c_i\zeta^j_b, \tag{III.6.22}$$

$$\Gamma^j_{ik}=\frac{h^{j\ell}}{2}\left(\frac{\partial h_{\ell i}}{\partial\xi^k}+\frac{\partial h_{k\ell}}{\partial\xi^i}-\frac{\partial h_{ik}}{\partial\xi^\ell}\right). \tag{III.6.22a}$$

For the case of σ a surface of constant time, (III.6.21) reduces to (III.2.18).

So, *if the initial and final "times" are generalized to arbitrary spacelike surfaces, the conditions* (III.6.21) *obeyed by the initial and final wave functionals do not truly break Lorentz invariance,* since *they depend only on the surfaces on which the states are attached, and not on any absolute time direction.* Thus *the entire theory can be formulated in a manifestly Lorentz invariant way,* at least formally. With arbitrary initial and final surfaces, any lattice realization of the path integrals in (III.6.3) will in general involve a non-Cartesian lattice. There are doubtless difficulties in defining such integrals, but they are beyond the scope of the present work.

III.7 Conclusions

In this chapter, I have developed and examined the sum-over-histories formulation of generalized quantum mechanics for a non-Abelian gauge theory in the absence of matter, which in addition to its inherent interest can be viewed as a toy model for Einstein's general relativity. The path integrals have been explicitly defined via an infinitesimal lattice, and shown to be gauge invariant.

The most general form of the theory allows any set of gauge invariant phase space alternatives to be assigned a decoherence functional. Restricting the alternatives to the phase space implementations of the gauge electric and magnetic fields and the covariant derivative gives the "physical phase space formulation". If instead only gauge invariant configuration space alternatives are considered, we obtain a different subset of possible coarse grainings. This theory is formally Lorentz-invariant as well. A further restriction to coarse grainings involving the configuration space implementations of gauge electric and magnetic fields and covariant derivative gives the "physical configuration space formulation".

We have shown that the physical phase space formulation agrees with a reduced phase space canonical operator (or, as it is known in other works including [24], "ADM") formulation, so long as the coarse grainings did not involve time derivatives. In particular, the non-Abelian Gauss's law constraint $\mathbf{D} \cdot \boldsymbol{\pi} = 0$ is always satisfied.

The physical configuration space formulation behaves slightly differently. One formally defined quantity which roughly corresponds to the longitudinal electric field E_L from E&M was shown to behave in the same way as E_L did in the Abelian theory. *I.e.*, coarse grainings by this quantity which decohere predict that it vanishes. However, if one coarse grains by more complicated quantities related to the configuration space constraint $-\mathbf{D} \cdot \mathbf{E}$, that may not be so. In E&M, we have explicitly shown that for suitable initial conditions, coarse grainings by one such quantity $\{g[\mathbf{B}(t_i)] \left|\langle E_L \rangle\right|^2\}$ decohere and predict non-zero probabilities for the quantity not to vanish.

Despite the disagreement between the physical configuration space implementation and reduced phase space operator quantization, the sum-over-histories formulation is still attractive, since it can be expressed in a manifestly Lorentz-invariant form. On the other hand, the operator theory gives special consideration to the time direction by singling out the constraint, which is just the time component of the equations of motion $D_\mu G^{\mu\nu} = 0$, to be identically satisfied.

Since the disagreement between the sum-over-histories theory and a natural extension of the operator theory comes about when the coarse graining involves quantities averaged over a spacetime region, as opposed to the usual quantum mechanical expressions involving alternatives defined at a single moment of time, perhaps the sum-over-histories and reduced phase space methods should be seen as different generalizations of the previously tested formulations (in which the quantity considered here is not accessible). The Lorentz invariance of the sum-over-histories method then makes it the preferred generalization in light of the potential application to quantum gravity, as it takes one more step towards eliminating the special role of time in the theory.

There is also some question as to whether one could construct a physical apparatus to measure the involved quantity by which we coarse grained in section III.5.4; on a practical level, the fields are not directly measurable, but only accessible through their interactions with charged particles. It is conceivable that the differences between the sum-over-histories and operator formalisms are undetectable in their application to QED and QCD. However, it is reasonable to expect that the issues raised by the discrepancy between them will be relevant to a quantization of GR. Is enforcement of the constraints more fundamental than manifest diffeomorphism (here Lorentz) invariance, or should we only expect the constraints to be satisfied when the class of alternatives considered singles out the corresponding time direction in its choice of surfaces?

III.A Appendix: Calculation of class operator for section III.5.3

In this appendix we calculate

$$\mathcal{C}_\Delta = \left\{ \prod_{M=J}^{0} \int \mathcal{D}\mathcal{E}^M \exp\left[i\delta t \int d^3x \frac{1}{2}\left(\mathcal{E}^M\right)^2 \right] \right\} e_\Delta[\mathcal{E}], \tag{III.5.24}$$

where

$$e_\Delta[\mathcal{E}] = \int_\Delta d^{2n} f \delta(f - \langle\mathcal{E}\rangle) \tag{III.5.26}$$

describes the coarse graining by values of the mode average

$$\langle\mathcal{E}\rangle = \frac{1}{\Omega} \int_\Omega dt\, d^3k\, \mathcal{E}_{\mathbf{k}}(t) = \frac{1}{\Omega} \sum_{M\in\Omega} \delta t \int_\Omega d^3k\, \mathcal{E}^M_{\mathbf{k}}. \tag{III.5.27}$$

While $\mathcal{E}(\mathbf{x})$ is a real quantity, the Fourier transform

$$\mathcal{E}^M_{\mathbf{k}} = \int \frac{d^3x}{(2\pi)^{3/2}} e^{-i\mathbf{k}\cdot\mathbf{x}} \mathcal{E}^M(\mathbf{x}) \tag{III.5.28}$$

is complex but constrained to obey $\mathcal{E}^*_{\mathbf{k}} = \mathcal{E}_{-\mathbf{k}}$. Integrating over the independent degrees of freedom in Fourier space necessitates the development of more notation. Letting a superscript of R or I indicate the real or imaginary part, respectively, of a complex number, and using the Jacobian determinant calculated in appendix III.B for the discrete Fourier transform, the path integral measure is (using the infinite numerical constant Ξ defined in appendix III.B)

$$\begin{aligned} \mathcal{D}\mathcal{E}^M &= \prod_{a,\mathbf{x}} N_A d\mathcal{E}^M_a(\mathbf{x}) = \prod_{a,\mathbf{x}} N_A d\mathcal{E}^{M\mathrm{R}}_a(\mathbf{x}) d\mathcal{E}^{M\mathrm{I}}_a(\mathbf{x}) \delta\big(\mathcal{E}^{M\mathrm{I}}_a(\mathbf{x})\big) \\ &= \prod_a \left(\prod_{\mathbf{k}} N_A \frac{\delta^3 k}{\delta^3 x} d\mathcal{E}^{M\mathrm{R}}_{a,\mathbf{k}} d\mathcal{E}^{M\mathrm{I}}_{a,\mathbf{k}} \right) \Xi \prod_{\mathbf{k}}^{1/2} \frac{\delta^3 x}{\delta^3 k} \delta\left(\mathcal{E}^{M\mathrm{R}}_{a,\mathbf{k}} - \mathcal{E}^{M\mathrm{R}}_{a,-\mathbf{k}}\right) \delta\left(\mathcal{E}^{M\mathrm{I}}_{a,\mathbf{k}} + \mathcal{E}^{M\mathrm{I}}_{a,-\mathbf{k}}\right), \end{aligned} \tag{III.A.1}$$

where the $\prod^{1/2}$ means we are only taking the product over half the modes (leaving out the redundant ones, whose spatial frequency is minus the spatial frequency of a mode already counted), so that $\prod_{\mathbf{x}} f^{\mathrm{R}}(\mathbf{x})$ and $\prod_{\mathbf{k}}^{1/2} f^{\mathrm{R}}_{\mathbf{k}} f^{\mathrm{I}}_{\mathbf{k}}$ each have the same number of factors. The factor in $\prod^{1/2}$ for the zero mode[19] is understood to be

$$\left(\frac{\delta^3 x}{\delta^3 k}\right)^{1/2} \delta(\mathcal{E}^{M\mathrm{I}}_{a,\mathbf{0}}). \tag{III.A.1a}$$

We can use the delta functions to perform the integrals over half of the Fourier components so that

$$\mathcal{D}\mathcal{E}^M = \prod_a \Xi \prod_{\mathbf{k}}^{1/2} N_A^2 \frac{\delta^3 k}{\delta^3 x} d\mathcal{E}^{M\mathrm{R}}_{a,\mathbf{k}} d\mathcal{E}^{M\mathrm{I}}_{a,\mathbf{k}} = \Xi^n \prod_{\mathbf{k}}^{1/2} \left(N_A^2 \frac{\delta^3 k}{\delta^3 x}\right)^n d^n\mathcal{E}^{M\mathrm{R}}_{\mathbf{k}} d^n\mathcal{E}^{M\mathrm{I}}_{\mathbf{k}} \tag{III.A.2}$$

with the factor for the zero mode understood to be

$$\left(N_A^2 \frac{\delta^3 k}{\delta^3 x}\right)^{n/2} d^n\mathcal{E}^{M\mathrm{R}}_{\mathbf{0}}. \tag{III.A.2a}$$

The relevant part of the Lagrangian is

$$\frac{1}{2}\int d^3x \left[\mathcal{E}^M(\mathbf{x})\right]^2 = \frac{1}{2}\int d^3x \left|\mathcal{E}^M(\mathbf{x})\right|^2 = \frac{1}{2}\int d^3k \left|\mathcal{E}^M_{\mathbf{k}}\right|^2 = \sum_{\mathbf{k}}^{1/2} \delta^3 k \left|\mathcal{E}^M_{\mathbf{k}}\right|^2; \tag{III.A.3}$$

defining $\omega = \delta t \delta^3 k$, we have

$$\mathcal{C}_\Delta = \left(\prod_{M=J}^{0} \Xi^n \prod_{\mathbf{k}}^{1/2} \left(N_A^2 \frac{\delta^3 k}{\delta^3 x}\right)^n \int d^n\mathcal{E}^{M\mathrm{R}}_{\mathbf{k}} d^n\mathcal{E}^{M\mathrm{I}}_{\mathbf{k}} \exp\left\{i\omega\left[\left(\mathcal{E}^{M\mathrm{R}}_{\mathbf{k}}\right)^2 + \left(\mathcal{E}^{M\mathrm{I}}_{\mathbf{k}}\right)^2\right]\right\}\right) e_\Delta[\mathcal{E}]. \tag{III.A.4}$$

If we use λ as a mode label, combining M and $\mathbf{k}$,

$$e_\Delta[\mathcal{E}] = \int_\Delta d^n f^{\mathrm{R}} d^n f^{\mathrm{I}} \delta^n\Big(f^{\mathrm{R}} - \frac{\omega}{\Omega}\sum_{\lambda\in\Omega}\mathcal{E}^{\mathrm{R}}_\lambda\Big)\delta^n\Big(f^{\mathrm{I}} - \frac{\omega}{\Omega}\sum_{\lambda\in\Omega}\mathcal{E}^{\mathrm{I}}_\lambda\Big) = \int_\Delta d^{2n} f \delta^{2n}\Big(f - \frac{\omega}{\Omega}\sum_{\lambda\in\Omega}\mathcal{E}_\lambda\Big) \tag{III.A.5}$$

and

$$\mathcal{C}_\Delta = \left(\Xi^{n(J+1)} \prod_\lambda \left(N_A^2 \frac{\delta^3 x}{\delta^3 k}\right)^n \int d^n\mathcal{E}^{\mathrm{R}}_\lambda d^n\mathcal{E}^{\mathrm{I}}_\lambda \exp\left\{i\omega\left[\left(\mathcal{E}^{\mathrm{R}}_\lambda\right)^2 + \left(\mathcal{E}^{\mathrm{I}}_\lambda\right)^2\right]\right\}\right) e_\Delta[\mathcal{E}]. \tag{III.A.6}$$

We can factor the product in (III.A.6) into a product over modes in Ω and one over modes not in Ω. The latter is a constant which is the same for all alternatives $\{c_\Delta\}$:

$$\mathcal{K} = \left(\Xi^{n(J+1)} \prod_{\lambda\notin\Omega} \left(N_A^2 \frac{\delta^3 x}{\delta^3 k}\right)^n \int d^n\mathcal{E}^{\mathrm{R}}_\lambda d^n\mathcal{E}^{\mathrm{I}}_\lambda \exp\left\{i\omega\left[\left(\mathcal{E}^{\mathrm{R}}_\lambda\right)^2 + \left(\mathcal{E}^{\mathrm{I}}_\lambda\right)^2\right]\right\}\right), \tag{III.A.7}$$

[19]There will typically also be modes on the boundary of spatial frequency space which are identified with the corresponding modes on the opposite boundary, and so that for these $\mathbf{k}$'s $\mathcal{E}_{\mathbf{k}} = \mathcal{E}_{-\mathbf{k}}$ as with the zero mode $\mathbf{k} = \mathbf{0}$. For example, in the discrete Fourier transform on a one-dimensional lattice [47] with an even number N of points, the modes of frequency $1/2N$ and $-1/2N$ (the Nyquist critical frequency and its image) are identified, so the situation is analogous to that of the zero mode. The identification, when combined with the condition $\mathcal{E}^*_{\mathbf{k}} = \mathcal{E}_{-\mathbf{k}}$, requires that the Fourier components on the boundary be real. The boundary is not a region of interest to us in spatial frequency space, and we assume that the prescription for those factors is similar to the one for the zero mode.

which leaves

$$\begin{aligned}\mathcal{C}_\Delta =&\mathcal{K}\Bigg(\prod_{\lambda\in\Omega}\left(N_A^2\frac{\delta^3x}{\delta^3k}\right)^n\int d^n\mathcal{E}_\lambda^{\mathrm{R}}d^n\mathcal{E}_\lambda^{\mathrm{I}}\exp\left\{i\omega\left[\left(\mathcal{E}_\lambda^{\mathrm{R}}\right)^2+\left(\mathcal{E}_\lambda^{\mathrm{I}}\right)^2\right]\right\}\Bigg)\\ &\times\int_\Delta d^{2n}f\,\delta^{2n}\Bigg(f-\frac{\omega}{\Omega}\sum_{\lambda\in\Omega}\mathcal{E}_\lambda\Bigg).\end{aligned}\tag{III.A.8}$$

If we define $N=\sum_{\lambda\in\Omega}1$ to be the number of modes in Ω, and write all the modes of $\mathcal{E}$ in Ω as a $2N$-component column vector:

$$\mathcal{E}=\begin{pmatrix}\mathcal{E}_\lambda^{\mathrm{R}}\\ \mathcal{E}_\lambda^{\mathrm{I}}\end{pmatrix},\tag{III.A.9}$$

we have

$$\mathcal{C}_\Delta=\mathcal{K}\left[\prod_{\lambda\in\Omega}\left(N_A^2\frac{\delta^3x}{\delta^3k}\right)^n\int d^n\mathcal{E}_\lambda^{\mathrm{R}}d^n\mathcal{E}_\lambda^{\mathrm{I}}\right]e^{i\omega\mathcal{E}^2}\int_\Delta d^{2n}f\,\delta^{2n}\Bigg(f-\frac{\omega}{\Omega}\sum_{\lambda\in\Omega}\mathcal{E}_\lambda\Bigg).\tag{III.A.10}$$

We define a column vector Υ to be the discrete Fourier transform of $\mathcal{E}$:

$$\Upsilon=\begin{pmatrix}\Upsilon_{w\mathbf{y}}^{\mathrm{R}}\\ \Upsilon_{w\mathbf{y}}^{\mathrm{I}}\end{pmatrix}=\begin{pmatrix}\frac{1}{2}&\frac{1}{2}\\ \frac{1}{2i}&-\frac{1}{2i}\end{pmatrix}\begin{pmatrix}\frac{\omega}{\Omega}e^{i\mathbf{y}\cdot\mathbf{k}+iwt}&0\\ 0&\frac{\omega}{\Omega}e^{-i\mathbf{y}\cdot\mathbf{k}-iwt}\end{pmatrix}\begin{pmatrix}1&i\\ 1&-i\end{pmatrix}\begin{pmatrix}\mathcal{E}_{t\mathbf{k}}^{\mathrm{R}}\\ \mathcal{E}_{t\mathbf{k}}^{\mathrm{I}}\end{pmatrix}=M\mathcal{E}.\tag{III.A.11}$$

[This is a rigorous version of the traditional treatment of the complex Υ and Υ^* as independent variables; the middle matrix of the product of three is the one which would be used to convert the column vector $\begin{pmatrix}\mathcal{E}_{t\mathbf{k}}\\ \mathcal{E}_{t\mathbf{k}}^*\end{pmatrix}$ to $\begin{pmatrix}\Upsilon_{w\mathbf{y}}\\ \Upsilon_{w\mathbf{y}}^*\end{pmatrix}$.] The zero components of Υ are

$$\Upsilon_{0\mathbf{0}}=\frac{\omega}{\Omega}\sum_{\lambda\in\Omega}\mathcal{E}_\lambda=\langle\mathcal{E}\rangle.\tag{III.A.12}$$

Since M is a real matrix which is the product of three matrices, each of which is proportional to a unitary matrix, it must be proportional to an orthogonal (i.e., real and unitary) matrix $\mathcal{M}$. The calculation in appendix III.B yields

$$\det M=N^{-N}=(\det\mathcal{M})(N^{-1/2})^{2N},\tag{III.A.13}$$

so we have $M=\mathcal{M}/\sqrt{N}$. Thus

$$\Upsilon^2=\mathcal{E}^{\mathrm{TR}}M^{\mathrm{TR}}M\mathcal{E}=\frac{\mathcal{E}^{\mathrm{TR}}\mathcal{M}^{\mathrm{TR}}\mathcal{M}\mathcal{E}}{N}=\frac{\mathcal{E}^2}{N},\tag{III.A.14}$$

so $\omega\mathcal{E}^2 = N\omega\Upsilon^2 = \Omega\Upsilon^2$ and

$$\begin{aligned}\mathcal{C}_\Delta =&\mathcal{K}\left[\prod_{w\mathbf{y}}\left(N_A^2\frac{\delta^3 x}{\delta^3 k}\right)^n \int N d^n\Upsilon^{\mathrm{R}}_{w\mathbf{y}} d^n\Upsilon^{\mathrm{I}}_{w\mathbf{y}}\right] e^{i\Omega\Upsilon^2}\int\limits_\Delta d^{2n}f\,\delta^{2n}(f-\Upsilon_{\mathbf{00}})\\ =&\mathcal{K}\Bigg(\prod_{w\mathbf{y}}\left(N_A^2\frac{\delta^3 x}{\delta^3 k}\right)^n \int N d^n\Upsilon^{\mathrm{R}}_{w\mathbf{y}} d^n\Upsilon^{\mathrm{I}}_{w\mathbf{y}} \exp\left\{i\Omega\left[\left(\Upsilon^{\mathrm{R}}_{w\mathbf{y}}\right)^2+\left(\Upsilon^{\mathrm{I}}_{w\mathbf{y}}\right)^2\right]\right\}\Bigg)\\ &\times\int\limits_\Delta d^{2n}f\,\delta^{2n}(f-\Upsilon_{\mathbf{00}}).\end{aligned} \tag{III.A.15}$$

Factoring all the $\{\Upsilon_{w\mathbf{y}}\}$ except the zero mode into the constant, and using the delta function to do the $\Upsilon_{\mathbf{00}}$ integrals, we have

$$\mathcal{C}_\Delta = \mathcal{K}'\int\limits_\Delta d^{2n}f\, e^{i\Omega|f|^2}. \tag{III.A.16}$$

III.B Appendix: Calculation of Jacobian determinants for discrete Fourier transforms

Given a complex function $f(x)$ of a D dimensional variable $x=\{x_\alpha|\alpha=1,2,\dots D\}$, if we define f only on a spatial lattice with N_α lattice points in the α direction (and thus $\prod_{\alpha=1}^D N_\alpha \equiv N$ total lattice points), we have a vector

$$f=\begin{pmatrix} f_x^{\mathrm{R}}\\ f_x^{\mathrm{I}}\end{pmatrix} \tag{III.B.1}$$

with $2N$ real components. If we take the Fourier transform (see [47] for a general treatment of the discrete Fourier transform)

$$F_k=\sum_x e^{\mp ik\cdot x} f_x, \tag{III.B.2}$$

there is a corresponding matrix transformation[20] on $\mathbb{R}^{2N}$:

$$\begin{pmatrix} F_k^{\mathrm{R}}\\ F_k^{\mathrm{I}}\end{pmatrix}\begin{pmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2i} & -\frac{1}{2i}\end{pmatrix}\begin{pmatrix}\exp(\mp i\sum_\alpha k_\alpha x_\alpha) & 0\\ 0 & \exp(\pm i\sum_\alpha k_\alpha x_\alpha)\end{pmatrix}\begin{pmatrix}1 & i\\ 1 & -i\end{pmatrix}\begin{pmatrix} f_x^{\mathrm{R}}\\ f_x^{\mathrm{I}}\end{pmatrix} \tag{III.B.3}$$

or

$$F=Mf. \tag{III.B.3$'$}$$

The Jacobian of this transformation is given by $\det M$. Since the first and third of the three matrices of which M is a product are inverses of each other and the second is block diagonal,

[20] As with (III.A.11), this treatment is the more careful analog of treating f_x and f_x^* as independent variables.

we have

$$\det_{2N\times 2N} M = \det_{N\times N} e^{\mp ik\cdot x} \det_{N\times N} e^{\pm ik\cdot x} = \det_{N\times N}\left(\sum_k e^{\mp ix\cdot k} e^{\pm ik\cdot y}\right) = \det_{N\times N}(N\delta_{xy}) = N^N. \tag{III.B.4}$$

To apply this to the transformations in section III.5.3, we need to take into account the normalization constants. In the discrete case, (III.5.28) becomes

$$\mathcal{E}^M_{a,\mathbf{k}} = \sum_{\mathbf{x}} \frac{\delta^3 x}{(2\pi)^{3/2}} e^{-i\mathbf{k}\cdot\mathbf{x}} \mathcal{E}^M_a(\mathbf{x}) \tag{III.B.5}$$

and the Jacobian is

$$N^N \left(\frac{\delta^3 x}{(2\pi)^{3/2}}\right)^{2N}. \tag{III.B.6}$$

Now, the relationship between $\delta^3 x$, the number of spatial lattice sites N, and the lattice spacing $\delta^3 k$ in spatial frequency can be deduced by geometric arguments, but the simplest method is to note that

$$\mathcal{E}^M_a(\mathbf{x}) = \sum_{\mathbf{k}} \frac{\delta^3 k}{(2\pi)^{3/2}} e^{i\mathbf{x}\cdot\mathbf{k}} \mathcal{E}^M_{a,\mathbf{k}} \tag{III.B.7}$$

and hence

$$\begin{pmatrix} \mathcal{E}^{M\mathrm{R}}_a(\mathbf{x}) \\ \mathcal{E}^{M\mathrm{I}}_a(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2i} & -\frac{1}{2i} \end{pmatrix} \begin{pmatrix} \frac{\delta^3 k}{(2\pi)^{3/2}} e^{i\mathbf{x}\cdot\mathbf{k}} & 0 \\ 0 & \frac{\delta^3 k}{(2\pi)^{3/2}} e^{-i\mathbf{x}\cdot\mathbf{k}} \end{pmatrix}$$
$$\times \begin{pmatrix} \frac{\delta^3 x}{(2\pi)^{3/2}} e^{-i\mathbf{k}\cdot\mathbf{y}} & 0 \\ 0 & \frac{\delta^3 x}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{y}} \end{pmatrix} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \begin{pmatrix} \mathcal{E}^{M\mathrm{R}}_a(\mathbf{y}) \\ \mathcal{E}^{M\mathrm{I}}_a(\mathbf{y}) \end{pmatrix} \tag{III.B.8}$$

Taking the determinant, we find

$$1 = \left(\frac{\delta^3 k}{(2\pi)^{3/2}}\right)^{2N} N^N \left(\frac{\delta^3 x}{(2\pi)^{3/2}}\right)^{2N} N^N \tag{III.B.9}$$

or

$$N = \frac{(2\pi)^3}{\delta^3 k \delta^3 x}. \tag{III.B.10}$$

Substituting into (III.B.6), we see that the Jacobian is

$$\left(\frac{(2\pi)^3}{\delta^3 k \delta^3 x}\right)^N \left(\frac{\delta^3 x}{(2\pi)^{3/2}}\right)^{2N} = \left(\frac{\delta^3 x}{\delta^3 k}\right)^N \tag{III.B.11}$$

so that

$$\prod_{\mathbf{x}} d\mathcal{E}^{M\mathrm{R}}_a(\mathbf{x}) d\mathcal{E}^{M\mathrm{I}}_a(\mathbf{x}) = \left(\frac{\delta^3 k}{\delta^3 x}\right)^N \prod_{\mathbf{k}} d\mathcal{E}^{M\mathrm{R}}_{a,\mathbf{k}} d\mathcal{E}^{M\mathrm{I}}_{a,\mathbf{k}} = \prod_{\mathbf{k}} \left(\frac{\delta^3 k}{\delta^3 x}\right) d\mathcal{E}^{M\mathrm{R}}_{a,\mathbf{k}} d\mathcal{E}^{M\mathrm{I}}_{a,\mathbf{k}}, \tag{III.B.12}$$

which is the correct factor for (III.A.1).

Equation (III.A.1) also involves the Jacobian for the transformation of the delta functions

$$\prod_{\mathbf{x}} \delta\big(\mathcal{E}_a^{M\mathrm{I}}(\mathbf{x})\big) \tag{III.B.13}$$

into

$$\prod_{\mathbf{k}}^{1/2} \delta\left(\mathcal{E}_{a,\mathbf{k}}^{M\mathrm{R}} - \mathcal{E}_{a,-\mathbf{k}}^{M\mathrm{R}}\right) \delta\left(\mathcal{E}_{a,\mathbf{k}}^{M\mathrm{I}} + \mathcal{E}_{a,-\mathbf{k}}^{M\mathrm{I}}\right). \tag{III.B.14}$$

To determine that, define $F_k^{\pm} = F_k \pm F_{-k}$ and observe that

$$\begin{aligned}\prod_x df_x^{\mathrm{R}} df_x^{\mathrm{I}} &= N^N \prod_k dF_k^{\mathrm{R}} dF_k^{\mathrm{I}} = N^N \prod_k^{1/2} dF_k^{\mathrm{R}} dF_{-k}^{\mathrm{R}} dF_k^{\mathrm{I}} dF_{-k}^{\mathrm{I}} \\ &= N^N \prod_k^{1/2} \frac{dF_k^{\mathrm{R}+} dF_k^{\mathrm{R}-}}{2} \frac{dF_k^{\mathrm{I}+} dF_k^{\mathrm{I}-}}{2},\end{aligned} \tag{III.B.15}$$

so

$$\prod_x \delta(f_x^{\mathrm{R}})\delta(f_x^{\mathrm{I}}) = N^{-N} \prod_k^{1/2} 2\delta(F_k^{\mathrm{R}+})\delta(F_k^{\mathrm{R}-}) 2\delta(F_k^{\mathrm{I}+})\delta(F_k^{\mathrm{I}-}). \tag{III.B.16}$$

We assume by symmetry that when we factor the Jacobian splits evenly:

$$\prod_x \delta(f_x^{\mathrm{R}}) = N^{-N/2} \prod_k^{1/2} 2\delta(F_k^{\mathrm{R}+})\delta(F_k^{\mathrm{I}-}) \tag{III.B.17a}$$

$$\prod_x \delta(f_x^{\mathrm{I}}) = N^{-N/2} \prod_k^{1/2} 2\delta(F_k^{\mathrm{R}-})\delta(F_k^{\mathrm{I}+}); \tag{III.B.17b}$$

this means that

$$\prod_{\mathbf{x}} \delta\big(\mathcal{E}_a^{M\mathrm{I}}(\mathbf{x})\big) = \Xi \prod_{\mathbf{k}}^{1/2} \left(\frac{\delta^3 x}{\delta^3 k}\right) \delta\left(\mathcal{E}_{a,\mathbf{k}}^{M\mathrm{R}} - \mathcal{E}_{a,-\mathbf{k}}^{M\mathrm{R}}\right) \delta\left(\mathcal{E}_{a,\mathbf{k}}^{M\mathrm{I}} + \mathcal{E}_{a,-\mathbf{k}}^{M\mathrm{I}}\right). \tag{III.B.18}$$

Ξ is not quite equal to $2^{N/2}$ because the analysis above does not go through for the zero mode and some modes on the boundary (see footnote 19, page 100) which are identified with their images. In those cases, the analysis produces the same Jacobian, only without the factor of 2. At any rate, Ξ is a constant, and its precise value is unimportant.

The determinant of the transformation (III.A.11) is even more straightforward. There the number of modes is just $N = \frac{\Omega}{\omega}$, and the determinant is thus

$$\left(\frac{\omega}{\Omega}\right)^{2N} N^N = N^{-N}. \tag{III.B.19}$$

III.C Appendix: Calculation of the decoherence functional for section III.5.4 in the presence of a Gaussian initial state

Here we calculate the decoherence functional (III.5.48) for an initial state where $p(b)$ [*cf.* (III.5.46)] is a Gaussian in $B = \frac{1}{b}$:

$$p(B^{-1}) = A_\lambda \Theta(B-\lambda) e^{-B^2/2\sigma^2}. \tag{III.C.1}$$

Then (III.5.48a) becomes

$$G(y) \propto \int_\lambda^\infty dB\, e^{i\Omega y B} e^{-B^2/2\sigma^2} = \Lambda_1(y) + e^{-(\Omega\sigma y)^2/2} \int_0^\infty dB e^{-(B-i\Omega\sigma^2 y)^2/2\sigma^2}, \tag{III.C.2}$$

where

$$\Lambda_1(y) = -\int_0^\lambda dB\, e^{-B^2/2\sigma^2} e^{i\Omega y B} \tag{III.C.3}$$

satisfies

$$|\Lambda_1(y)| < \lambda. \tag{III.C.4}$$

The second term in (III.C.2) can be massaged by deformation of contour to give

$$G(y) \propto \sqrt{\frac{2}{\pi}} \frac{\Lambda_1(y)}{\sigma} + G_R(y) + iG_I(y), \tag{III.C.5}$$

where

$$G_R(y) = e^{-(\Omega\sigma y)^2/2} \tag{III.C.5a}$$

and

$$G_I(y) = \sqrt{\frac{2}{\pi}} \Omega\sigma \int_0^y dz\, e^{(\Omega\sigma)^2(z^2-y^2)/2}. \tag{III.C.5b}$$

If we choose the bins to be of a uniform size Δ:

$$\Delta_J \equiv [J\Delta, (J+1)\Delta), \quad 0 \le J \in \mathbb{Z}, \tag{III.C.6}$$

we have

$$\operatorname{Re} D(J,J') \propto \int_{J\Delta}^{(J+1)\Delta} df \int_{J'\Delta}^{(J'+1)\Delta} df' e^{-(\Omega\sigma)^2(f-f')^2/2} + \sqrt{\frac{2}{\pi}} \frac{\Lambda_2(J-J',\Delta)\Delta^2}{\sigma}, \tag{III.C.7}$$

where

$$\Lambda_2(J-J',\Delta) = \Delta^{-2} \int_{J\Delta}^{(J+1)\Delta} df \int_{J'\Delta}^{(J'+1)\Delta} df' \operatorname{Re} \Lambda_1(f-f') \tag{III.C.8}$$

again satisfies

$$|\Lambda_2(J-J',\Delta)| < \lambda. \tag{III.C.9}$$

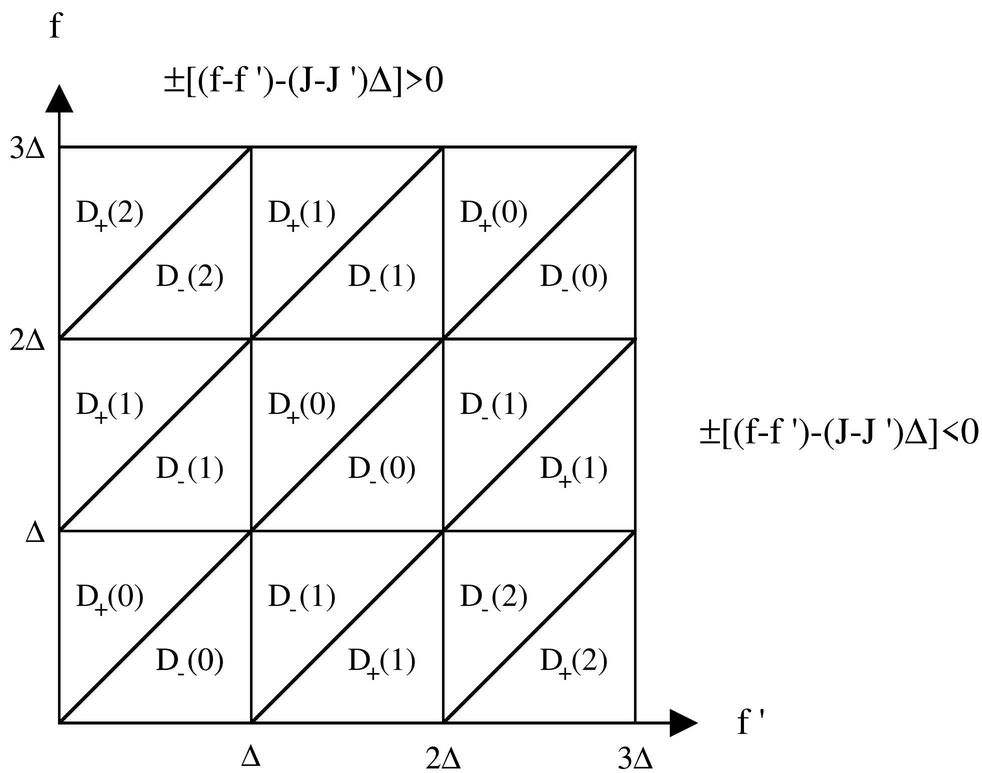


Figure III.2: The regions of integration for $D_{\pm}(|\Delta J|)$. Because of the exponential dropoff in $\mathrm{Re}G(f-f')$ as one moves towards larger $|f-f'|$, $D_+(|\Delta J|)$ is reduced from $D_+(0)$ by a factor of $e^{-(\Omega\sigma\Delta)^2(\Delta J)^2/2}$, and $D_-(|\Delta J|)$ is reduced from $D_-(1)$ by a factor of $e^{-(\Omega\sigma\Delta)^2(\Delta J-1)^2/2}$.

Making a suitable change of variables and factoring out Δ^2, we obtain

$$\mathrm{Re}\, D(J,J') \propto D_+(|J-J'|) + D_-(|J-J'|) + \sqrt{\frac{2}{\pi}}\frac{\Lambda_2(J-J',\Delta)}{\sigma}, \tag{III.C.10}$$

where $D_\pm$ is the contribution to the double integral from $\pm[f-f'-(J-J')\Delta](J-J')>0$:

$$D_\pm(\Delta J) = \int_0^1 d\eta\,(1-\eta)\exp[-(\Omega\sigma\Delta)^2(\Delta J \pm \eta)^2/2] \tag{III.C.10a}$$

(Fig. III.2). Now,

$$D_+(\Delta J) \le e^{-(\Omega\sigma\Delta)^2(\Delta J)^2/2}\int_0^1 d\eta\,(1-\eta)e^{-(\eta\Omega\sigma\Delta)^2/2} = e^{-(\Omega\sigma\Delta)^2(\Delta J)^2/2}D_+(0). \tag{III.C.11}$$

From the definition (III.C.10a) it is evident that $D_-(0) = D_+(0)$. For $\Delta J \geq 1$,

$$D_-(\Delta J) = \int_0^1 d\eta\, \eta \exp\{-(\Omega\sigma\Delta)^2[(\Delta J - 1) + \eta]^2/2\} \leq e^{-(\Omega\sigma\Delta)^2(\Delta J-1)^2/2} D_-(1). \quad \text{(III.C.12)}$$

Combining these results, we see[21]

$$\frac{\operatorname{Re} D(J+\Delta J, J)}{D(J,J)} \leq \frac{e^{-(\Omega\sigma\Delta)^2|\Delta J|^2/2} D_+(0) + e^{-(\Omega\sigma\Delta)^2(|\Delta J|-1)^2/2} D_-(1) + \sqrt{2/\pi}\, \Lambda_2(\Delta J, \Delta)/\sigma}{2D_+(0) + \sqrt{2/\pi}\, \Lambda_2(\Delta J, \Delta)/\sigma}. \quad \text{(III.C.13)}$$

So we have reduced the question of whether we have weak decoherence [$|\operatorname{Re} D(J+\Delta J)| \ll D(J,J)$ for $\Delta J \neq 0$] to a calculation of $D_+(0)$ and $D_-(1)$. It is straightforward to show

$$D_-(1) = \frac{1 - e^{-(\Omega\sigma\Delta)^2/2}}{(\Omega\sigma\Delta)^2} \quad \text{(III.C.14)}$$

and

$$D_+(0) = \frac{\sqrt{\pi/2}}{\Omega\sigma\Delta} \operatorname{erf}\left(\frac{\Omega\sigma\Delta}{\sqrt{2}}\right) - D_-(1), \quad \text{(III.C.15)}$$

where $\operatorname{erf} z$ is the error function $\operatorname{erf} z = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$, which satisfies $\operatorname{erf}(\infty) = 1$ and $\operatorname{erf} z \geq 1 - e^{-z^2}$. Thus

$$D_+(0) \geq \sqrt{\frac{\pi}{2}} \frac{1 - e^{-(\Omega\sigma\Delta)^2/2}}{\Omega\sigma\Delta} - \frac{1 - e^{-(\Omega\sigma\Delta)^2/2}}{(\Omega\sigma\Delta)^2} = \left(1 - e^{-(\Omega\sigma\Delta)^2/2}\right) \frac{\Omega\sigma\Delta\sqrt{\pi/2} - 1}{(\Omega\sigma\Delta)^2}. \quad \text{(III.C.16)}$$

This means that if the cutoff $\lambda \lesssim \sigma e^{-(\Omega\sigma\Delta)^2/2}$, (III.C.13) becomes, to lowest order in $e^{-(\Omega\sigma\Delta)^2/2}$,

$$\begin{aligned} \frac{\operatorname{Re} D(J+\Delta J, J)}{D(J,J)} &\lesssim \frac{\exp[-(\Omega\sigma\Delta)^2(|\Delta J| - 1)^2/2] D_-(1)}{2D_+(0)} + \mathcal{O}\left(\frac{\lambda}{\sigma}\right) \\ &\lesssim \frac{\exp[-(\Omega\sigma\Delta)^2(|\Delta J| - 1)^2/2]}{\Omega\sigma\Delta\sqrt{2\pi}} + \mathcal{O}\left(\frac{\lambda}{\sigma}\right). \end{aligned} \quad \text{(III.C.17)}$$

For large $\Omega\sigma\Delta$, the λ-dependent term will not be relevant to the issue of decoherence.

[21] Recall that $D(J,J)$ is real and positive by (I.3.4-I.3.4b).

Chapter IV

Modelling the Decoherence of Spacetime

IV.1 Introduction

In this chapter, I examine a phenomenon known as the decoherence of spacetime, by which coarse grainings by long-wavelength features of the gravitational field may be made to decohere by tracing over the short-wavelength modes of the field. This differs from previous work[26, 27] which used an additional field to obtain decoherence of the gravitational field in cosmological models, in that the decoherence examined here is be induced in the gravitational field itself with no external matter field.

This work takes a perturbative approach to the actual quantization of the gravitational field, as described in Sec. IV.3. The field in the sum over histories is expressed as a background field which solves the Einstein equations, plus a small perturbation. The expansion of the gravitational action in powers of the perturbation gives a second order wave term, followed by a third order interaction term containing two derivatives. This chapter performs calculations on the toy model described in Sec. IV.3.3, where the tensor gravitational field is replaced by a scalar field with a similar action.

Section IV.4 demonstrates the effects of splitting the scalar field into long-wavelength and short-wavelength parts. The second-order wave propagation terms do not couple the long-wavelength modes (LWMs) to the short-wavelength modes (SWMs), but the third order interaction terms do, and can be classified by the number of SWM factors: zero, one, two or three. Temporarily removing the terms with one and three SWM factors leaves an action whose terms are all quadratic in the SWMs, or independent of them. Thus the trace over the SWMs can be performed explicitly, and this is done in Sec. IV.5. As described in Sec. IV.5.4, the perturbative corrections to the decoherence functional can cause elements which are finite in the non-interacting theory to vanish if the SWMs are in a thermal state whose temperature is sufficiently high. Then certain terms in the perturbation series can become large in the high-temperature limit, producing seemingly non-perturbative effects.

In Sec. IV.6 I demonstrate that reïnserting the terms with one and three SWM factors into the action has no substantial effect on the result of Sec. IV.5. The terms linear in the SWMs can be removed by completion of the square to recover the original result. The terms cubic in the SWMs are examined in a perturbation series, and each term is seen to be perturbatively finite, even in the high-temperature limit. So, according to the perturbative analysis, the effect of the cubic terms is to multiply the decoherence functional by a factor of order unity.

Section IV.7 applies the properties of the decoherence functional found in Sec. IV.5 to a class of practical coarse grainings, and describes some circumstances under which decoherence can be expected.

IV.2 Environment-induced decoherence and the influence phase

Decoherence in most physical systems is caused by a division into the "system" of interest, and an "environment" about which no information is gathered. In the language of generalized quantum mechanics, this means that the coarse graining is described by alternatives which refer only to the system variables. (See [12] for further details and a bibliography of prior work)

To describe this mathematically, we work with the decoherence functional (I.3.15) specialized to a suitably normalized initial state ρ and a condition of future indifference to give

$$D[\varphi_1,\varphi_2]=\rho[\varphi_1',\varphi_2']\delta[\varphi_2''-\varphi_1'']e^{i(S[\varphi_1]-S[\varphi_2])} \tag{IV.2.1}$$

If we make a division of φ into system variables Φ and environment variables ϕ, split up the action into a ϕ-independent piece $S_\Phi[\Phi]=S|_{\phi=0}$ and a piece S_E describing the environment and its interaction with the system:

$$S[\varphi]=S_\Phi[\Phi]+S_E[\phi,\Phi] \tag{IV.2.2}$$

and assume that the initial state is the product of uncorrelated states for the system and the environment:

$$\rho[\varphi]=\rho_\Phi[\Phi_1',\Phi_2']\rho_\phi[\phi_1',\phi_2'] \tag{IV.2.3}$$

then the decoherence functional for a coarse-graining which makes no reference to the environment variables (but is still fine-grained in the system variables) can be written

$$D[\Phi_1,\Phi_2]=\int\mathcal{D}\phi_1\mathcal{D}\phi_2 D[\varphi_1,\varphi_2]=\rho_\Phi[\Phi_1',\Phi_2']\delta[\Phi_2''-\Phi_1'']e^{i(S_\Phi[\Phi_1]-S_\Phi[\Phi_2]+W[\Phi_1,\Phi_2])} \tag{IV.2.4a}$$

where

$$e^{iW[\Phi_1,\Phi_2]}=\int\mathcal{D}\phi_1\mathcal{D}\phi_2\rho_\phi[\phi_1',\phi_2']\delta[\phi_2''-\phi_1'']e^{i(S_E[\phi_1,\Phi_1]-S_E[\phi_2,\Phi_2])}. \tag{IV.2.4b}$$

$W[\Phi_1,\Phi_2]$ is called the Feynman-Vernon influence phase[48]; if the *influence functional* e^{iW} becomes small for $\Phi_1\neq\Phi_2$, the "off-diagonal" parts of $D[\Phi_1,\Phi_2]$ will be suppressed, causing alternatives defined in terms of Φ to decohere [12].

IV.3 Perturbative GR and the scalar toy model

The goal of this chapter is to perform the division described in Sec. IV.2 on vacuum gravity. The idea behind this is that for coarse grainings which deal only with averages over sufficiently large regions of spacetime, gravity should behave classically, and thus such coarse grainings should decohere.

Following, once again, the argument of [24] stated in Sec. I.4.5, I will confine attention to a generalized quantum mechanics of general relativity. Problems still remain in the explicit realization of the formal decoherence functional defined by (I.4.11), however. Probably the most well-grounded way to provide explicit expressions would be to skeletonize the metric geometries using the Regge calculus [49]. For the purposes of this dissertation, however, I will take the simpler approach of expanding the metric about some fixed background, and taking the metric perturbation to be a tensor field defined on the background spacetime.

IV.3.1 The abstract index notation

It will be useful to adopt a notation popularized by Wald [14] in which latin indices $a, b, c, \ldots$ are so-called "abstract" spacetime indices which mark the positions of covariant or contravariant indices when a tensor (and only a tensor, so we do not write the Christoffel symbols as Γ^c_{ab}) is described in a particular coördinate system. For example, the tensor g_{ab} has the components $\{g_{\mu\nu}\}$. In this notation, the abstract indices are not to be thought of as taking particular values, but are part of the notation, like $\vec{a}$ or $\mathbf{a}$ for a three-vector.

IV.3.2 Perturbative GR

Given an arbitrary metric $g_{ab}(1)$ (not necessarily a solution to the Einstein equation), we can use a solution g_{ab} to the vacuum Einstein equation $R_{ab} = 0$ to define a family of metrics

$$g_{ab}(\lambda) = g_{ab} + \lambda\gamma_{ab}, \tag{IV.3.1}$$

where $\gamma_{ab} = g_{ab}(1) - g_{ab}$. We can expand quantities constructed from $g_{ab}(\lambda)$ [such as the Ricci tensor $R_{ab}(\lambda)$], which are functions of g_{ab} and γ_{ab}, in powers of λ. An expansion of the Einstein equation to first order in λ (see Sec. 7.5 of [14]) gives a wave equation for γ_{ab}. This corresponds to expanding the gravitational action $S_G = (16\pi G)^{-1} \int \sqrt{|g(\lambda)|}\, d^4x R(\lambda)$ to second order. To model the self-interaction of gravity, it is necessary to expand the action to third order in γ_{ab}. This is done in Appendix IV.A, and the result is

$$\begin{aligned} S = \frac{1}{16\pi G} \int \sqrt{|g|}\, d^4x \Bigg\{ & \lambda^2 \left[-\frac{1}{4}(\nabla_c\gamma_{ab})(\nabla^c\gamma^{ab}) + \frac{1}{2}\gamma^{ab}R_{acbd}\gamma^{cd} \right] \\ & \lambda^3\gamma^{ab} \left[\frac{1}{2}(\nabla_c\gamma^d_a)(\nabla^c\gamma_{bd}) + \frac{1}{4}(\nabla_a\gamma^d_c)(\nabla_b\gamma^c_d) + \frac{1}{2}(\nabla_c\gamma^d_a)(\nabla_d\gamma^c_b) - R_{acb}{}^d\gamma^{ce}\gamma_{de} \right] \\ & + \mathcal{O}(\lambda^4) \Bigg\}. \end{aligned} \tag{IV.A.41}$$

The terms on the first line describe a free wave equation for a tensor field, while those on the second line provide a self-interaction.

IV.3.3 The scalar toy model

In the spirit of Chapters II and III, I will consider a toy model which focuses on some of the features of the action (IV.A.41) and glosses over others. For one thing, I will assume that the background is (flat) Minkowski spacetime. In a cosmological scenario, this should be a reasonable assumption if the length scales on the problem do not approach the Hubble scale cH_0^{-1}. A more serious simplification is to discard the tensor information in (IV.A.41) and consider a scalar field φ with action

$$S = -\frac{1}{2}\int d^{D+1}x(1-\ell\varphi)(\partial^\mu\varphi)(\partial_\mu\varphi). \tag{IV.3.2}$$

This is the most general action analogous to (IV.A.41) which can be constructed from a scalar field on a flat background. In $3+1$ dimensions, $D=3$, the coupling constant ℓ has units of length, and is of the order of $\ell_p/4\sqrt{\pi}$. The combination $\ell\varphi$ is dimensionless and represents the metric perturbation γ_{ab}, so it should be small relative to unity for the perturbation theory approach to be valid.

IV.4 Dividing the modes

We want to make a division of the field φ appearing in the action

$$S[\varphi] = \int_{-T/2}^{T/2} dt L(t) \tag{IV.4.1a}$$

$$L(t) = \frac{1}{2}\int d^D x(1-\ell\varphi)[\dot{\varphi}^2 - (\boldsymbol{\nabla}\varphi)^2] \tag{IV.4.1b}$$

into long-wavelength modes (LWMs), labelled by Φ, to act as the "system" and short-wavelength modes (SWMs), labelled by ϕ, to act as the "environment". For reasons of mathematical convenience, I will first make this division only in the spatial directions. First I reëxpress the lagrangian in terms of the Fourier transform

$$\varphi_{\mathbf{k}}(t) = \int \frac{d^D x}{(2\pi)^{D/2}} e^{-i\mathbf{k}\cdot\mathbf{x}}\varphi(\mathbf{x},t); \qquad \varphi(\mathbf{x},t) = \int \frac{d^D x}{(2\pi)^{D/2}} e^{i\mathbf{x}\cdot\mathbf{k}}\varphi_{\mathbf{k}}(t) \tag{IV.4.2}$$

to get

$$\begin{aligned} L(t) = {} & \frac{1}{2}\int d^D k(|\dot{\varphi}_{\mathbf{k}}|^2 - \mathbf{k}^2|\varphi_{\mathbf{k}}|^2) \\ & -\frac{\ell}{2}\int \frac{d^D k_1 d^D k_2 d^D k_3}{(2\pi)^{D/2}}\delta^D(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)(\varphi_{\mathbf{k}_1}\dot{\varphi}_{\mathbf{k}_2}\dot{\varphi}_{\mathbf{k}_3} + \mathbf{k}_2\cdot\mathbf{k}_3\varphi_{\mathbf{k}_1}\varphi_{\mathbf{k}_2}\varphi_{\mathbf{k}_3}) \end{aligned} \tag{IV.4.3}$$

For calculational purposes, it is useful to treat the Fourier modes as being defined on a lattice in momentum space with lattice spacing δk, which will later be taken to zero. We choose a momentum scale k_c at which to divide the modes into LWMs $\{\Phi_{\mathbf{N}}\}$ with momenta $\mathbf{q_N} = \mathbf{N}\delta k$ satisfying $|\mathbf{q_N}| < k_c$ and SWMs $\{\phi_{\mathbf{M}}\}$ with momenta $\mathbf{k_M} = \mathbf{M}\delta k$ satisfying $|\mathbf{k_M}| > k_c$. The mode labels $\mathbf{M}$ and $\mathbf{N}$ are both vectors lying in $\mathbb{Z}^D$. If we let $\delta^D k = (\delta k)^D$ and define the long-wavelength sector $\mathcal{L} = \{\mathbf{N} \mid |\mathbf{q_N}| < k_c\}$ and the short-wavelength sector $\mathcal{S} = \{\mathbf{M} \mid |\mathbf{k_M}| > k_c\}$, we can approximate the Fourier transform as

$$\begin{aligned}\varphi_{\mathbf{k}}(t) &= \sum_{\mathbf{N}\in\mathcal{L}} \frac{\Phi_{\mathbf{N}}(t)\delta_{\mathbf{kq_N}}}{(\delta^D k)^{1/2}} + \sum_{\mathbf{M}\in\mathcal{S}} \frac{\phi_{\mathbf{M}}(t)\delta_{\mathbf{kk_M}}}{(\delta^D k)^{1/2}} \\ &= \sum_{\mathbf{N}\in\mathcal{L}} (\delta^D k)^{1/2}\Phi_{\mathbf{N}}(t)\delta^D(\mathbf{k}-\mathbf{q_N}) + \sum_{\mathbf{M}\in\mathcal{S}} (\delta^D k)^{1/2}\phi_{\mathbf{M}}(t)\delta^D(\mathbf{k}-\mathbf{k_M}).\end{aligned} \tag{IV.4.4}$$

The normalization is chosen so that the part of the action quadratic in φ becomes

$$\frac{1}{2}\int d^D x[\dot{\varphi}^2 - (\boldsymbol{\nabla}\varphi)^2] = \frac{1}{2}\sum_{\mathbf{N}\in\mathcal{L}}\left(\left|\dot{\Phi}_{\mathbf{N}}\right|^2 - \mathbf{q}_{\mathbf{N}}^2\left|\Phi_{\mathbf{N}}\right|^2\right) + \frac{1}{2}\sum_{\mathbf{M}\in\mathcal{S}}\left(\left|\dot{\phi}_{\mathbf{M}}\right|^2 - \mathbf{k}_{\mathbf{M}}^2\left|\phi_{\mathbf{M}}\right|^2\right). \tag{IV.4.5}$$

Taking into account the fact that $\varphi(\mathbf{x})$ is real, which means $\varphi_{-\mathbf{k}} = \varphi_{\mathbf{k}}^*$, or $\Phi_{-\mathbf{N}} = \Phi_{\mathbf{N}}^*$ and $\phi_{-\mathbf{M}} = \phi_{\mathbf{M}}^*$, we can write any expression using only half of the complex modes, which define the other half by complex conjugation. We define $\mathcal{L}/2$ and $\mathcal{S}/2$ as arbitrarily chosen halves of $\mathcal{L}$ and $\mathcal{S}$ so that $\{\Phi_{\mathbf{N}}|\mathbf{N}\in\mathcal{L}/2\}$ and $\{\phi_{\mathbf{M}}|\mathbf{M}\in\mathcal{S}/2\}$ between them define $\varphi_{\mathbf{k}}$. This makes the noninteracting ($\ell = 0$) action

$$\frac{1}{2}\int d^D x[\dot{\varphi}^2 - (\boldsymbol{\nabla}\varphi)^2] = \sum_{\mathbf{N}\in\mathcal{L}/2}\left(\left|\dot{\Phi}_{\mathbf{N}}\right|^2 - \mathbf{q}_{\mathbf{N}}^2\left|\Phi_{\mathbf{N}}\right|^2\right) + \sum_{\mathbf{M}\in\mathcal{S}/2}\left(\left|\dot{\phi}_{\mathbf{M}}\right|^2 - \mathbf{k}_{\mathbf{M}}^2\left|\phi_{\mathbf{M}}\right|^2\right). \tag{IV.4.6}$$

which is the action of a set of uncoupled harmonic oscillators. The interaction terms can be classified by the number of factors of the "environment" field ϕ to give[1]

$$L[\varphi] = L(\phi,\Phi) = L_\Phi(\Phi) + L_0(\phi) + \lambda L_\phi(\phi,\Phi) + \lambda L_{\phi\phi}(\phi,\Phi) + \lambda L_{\phi\phi\phi}(\phi), \tag{IV.4.7}$$

[1] I have written the lagrangians as functions, *e.g.*, $L_\Phi(\Phi)$ rather than functionals, since their arguments are now collections of modes like $\{\Phi_{\mathbf{N}}\}$ rather functions of position $\mathbf{x}$, like $\varphi(\mathbf{x})$.

where $\lambda = \frac{\ell(\delta^D k)^{1/2}}{(2\pi)^{D/2}}$ is the coupling constant, and

$$\begin{aligned} L_\Phi(\Phi) = &\sum_{\mathbf{N}\in\mathcal{L}/2} \left(\left|\dot{\Phi}_\mathbf{N}\right|^2 - \mathbf{q}_\mathbf{N}^2 \left|\Phi_\mathbf{N}\right|^2 \right) \\ &- \frac{\lambda}{2} \sum_{\mathbf{N}_1,\mathbf{N}_2,\mathbf{N}_3\in\mathcal{L}} \delta_{\mathbf{N}_1+\mathbf{N}_2+\mathbf{N}_3,\mathbf{0}} (\Phi_{\mathbf{N}_1}\dot{\Phi}_{\mathbf{N}_2}\dot{\Phi}_{\mathbf{N}_3} + \mathbf{q}_{\mathbf{N}_2}\cdot\mathbf{q}_{\mathbf{N}_3}\Phi_{\mathbf{N}_1}\Phi_{\mathbf{N}_2}\Phi_{\mathbf{N}_3}) \end{aligned} \tag{IV.4.8a}$$

$$L_0(\phi) = \sum_{\mathbf{M}\in\mathcal{S}/2} \left(\left|\dot{\phi}_\mathbf{M}\right|^2 - \mathbf{k}_\mathbf{M}^2 \left|\phi_\mathbf{M}\right|^2 \right) \tag{IV.4.8b}$$

$$\begin{aligned} L_\phi(\phi,\Phi) = &- \frac{1}{2} \sum_{\substack{\mathbf{M}\in\mathcal{S}\\ \mathbf{N}_1,\mathbf{N}_2\in\mathcal{L}}} \delta_{\mathbf{M}+\mathbf{N}_1+\mathbf{N}_2,\mathbf{0}} [\phi_\mathbf{M}\dot{\Phi}_{\mathbf{N}_1}\dot{\Phi}_{\mathbf{N}_2} + 2\dot{\phi}_\mathbf{M}\dot{\Phi}_{\mathbf{N}_1}\Phi_{\mathbf{N}_2} \\ &+ (2\mathbf{k}_\mathbf{M} + \mathbf{q}_{\mathbf{N}_1})\cdot\mathbf{q}_{\mathbf{N}_2}\phi_\mathbf{M}\Phi_{\mathbf{N}_1}\Phi_{\mathbf{N}_2}] \end{aligned} \tag{IV.4.8c}$$

$$\begin{aligned} L_{\phi\phi}(\phi,\Phi) = &- \frac{1}{2} \sum_{\substack{\mathbf{M}_1,\mathbf{M}_2\in\mathcal{S}\\ \mathbf{N}\in\mathcal{L}}} \delta_{\mathbf{M}_1+\mathbf{M}_2+\mathbf{N},\mathbf{0}} [\Phi_\mathbf{N}\dot{\phi}_{\mathbf{M}_1}\dot{\phi}_{\mathbf{M}_2} + 2\dot{\Phi}_\mathbf{N}\dot{\phi}_{\mathbf{M}_1}\phi_{\mathbf{M}_2} \\ &+ (2\mathbf{q}_\mathbf{N} + \mathbf{k}_{\mathbf{M}_1})\cdot\mathbf{k}_{\mathbf{M}_2}\Phi_\mathbf{N}\phi_{\mathbf{M}_1}\phi_{\mathbf{M}_2}] \end{aligned} \tag{IV.4.8d}$$

$$L_{\phi\phi\phi}(\phi) = - \frac{1}{2} \sum_{\mathbf{M}_1,\mathbf{M}_2,\mathbf{M}_3\in\mathcal{S}} \delta_{\mathbf{M}_1+\mathbf{M}_2+\mathbf{M}_3,\mathbf{0}} (\phi_{\mathbf{M}_1}\dot{\phi}_{\mathbf{M}_2}\dot{\phi}_{\mathbf{M}_3} + \mathbf{k}_{\mathbf{M}_2}\cdot\mathbf{k}_{\mathbf{M}_3}\phi_{\mathbf{M}_1}\phi_{\mathbf{M}_2}\phi_{\mathbf{M}_3}). \tag{IV.4.8e}$$

Under this division of the field φ, we can perform the division of the decoherence functional described by (IV.2.4), with

$$S_E(\phi,\Phi) = S_0[\phi] + \lambda S_\phi[\phi,\Phi] + \lambda S_{\phi\phi}[\phi,\Phi] + \lambda S_{\phi\phi\phi}[\phi], \tag{IV.4.9}$$

IV.5 The quadratic terms

Since the parts of the action defined in (IV.4.8b) and (IV.4.8d) are quadratic in ϕ, it would be possible to do the path integrals in (IV.2.4b) explicitly if the action S_E included only those terms. Thus we turn our attention for the time being to the modified influence functional

$$e^{iW_{0+\phi\phi}[\Phi_1,\Phi_2]} = \int \mathcal{D}\phi_1\mathcal{D}\phi_2\rho_\phi(\phi_1',\phi_2')\delta(\phi_2''-\phi_1'')e^{i(S_{0+\phi\phi}[\phi_1,\Phi_1]-S_{0+\phi\phi}[\phi_2,\Phi_2])}. \tag{IV.5.1}$$

IV.5.1 A vector expression

The lagrangian $L_{0+\phi\phi}$ can be written, using $\phi^*_{\mathbf{M}} = \phi_{-\mathbf{M}}$, in the suggestive form

$$\begin{aligned} L_{0+\phi\phi}(\phi,\Phi) =& L_0(\phi) + \lambda L_{\phi\phi}(\phi,\Phi) \\ =& \frac{1}{2}\sum_{\mathbf{M}_1,\mathbf{M}_2\in\mathcal{S}} \Bigg\{ \dot{\phi}^*_{\mathbf{M}_1}(\delta_{\mathbf{M}_1\mathbf{M}_2} - \lambda\Phi_{\mathbf{M}_1-\mathbf{M}_2})\dot{\phi}_{\mathbf{M}_2} - \frac{d}{dt}(\lambda\phi^*_{\mathbf{M}_1}\dot{\Phi}_{\mathbf{M}_1-\mathbf{M}_2}\phi_{\mathbf{M}_2}) \\ & - \phi^*_{\mathbf{M}_1}\left[\delta_{\mathbf{M}_1\mathbf{M}_2}\mathbf{k}^2_{\mathbf{M}_1} - \lambda(k^2_{\mathbf{M}_1\mathbf{M}_2}\Phi_{\mathbf{M}_1-\mathbf{M}_2} + \ddot{\Phi}_{\mathbf{M}_1-\mathbf{M}_2})\right]\phi_{\mathbf{M}_2}\Bigg\} \end{aligned} \tag{IV.5.2}$$

where

$$k^2_{\mathbf{M}_1\mathbf{M}_2} = -\mathbf{q}_{\mathbf{M}_1-\mathbf{M}_2}\cdot\mathbf{k}_{-\mathbf{M}_1} - \mathbf{k}_{-\mathbf{M}_1}\cdot\mathbf{k}_{\mathbf{M}_2} - \mathbf{k}_{\mathbf{M}_2}\cdot\mathbf{q}_{\mathbf{M}_1-\mathbf{M}_2} = \mathbf{k}^2_{\mathbf{M}_1} + \mathbf{k}^2_{\mathbf{M}_1} - \mathbf{k}_{\mathbf{M}_1}\cdot\mathbf{k}_{\mathbf{M}_2} \tag{IV.5.3}$$

and we have extended the definition of Φ slightly to include $\Phi_{\mathbf{N}} = 0$ when $\mathbf{N} \notin \mathcal{L}$. We would like to write (IV.5.2) as a matrix expression in terms of a vector which describes the short-wavelength modes $\{\phi_{\mathbf{M}}\}$. However, the reality conditions $\phi^{\mathrm{R}}_{-\mathbf{M}} = \phi^{\mathrm{R}}_{\mathbf{M}}$ and $\phi^{\mathrm{I}}_{-\mathbf{M}} = -\phi^{\mathrm{I}}_{\mathbf{M}}$, which make the measure for the integral over independent modes

$$\mathcal{D}\phi \propto \prod_{\mathbf{M}\in\mathcal{S}/2} \mathcal{D}\phi^{\mathrm{R}}_{\mathbf{M}}\mathcal{D}\phi^{\mathrm{I}}_{\mathbf{M}}, \tag{IV.5.4}$$

necessitate some caution. There are several possibilities to consider:

- a complex vector ϕ_{ex} with components $\{\phi_{\mathbf{M}}|\mathbf{M}\in\mathcal{S}\}$. (It is useful to define the space of such vectors as $\mathbb{C}^{\mathcal{S}}$.) This has the advantage that the lagrangian (IV.5.2) is easily written in that form,

  $$L_{0+\phi\phi} = \frac{1}{2}\left[\dot{\phi}^\dagger_{\mathrm{ex}} m_{\mathrm{ex}}\dot{\phi}_{\mathrm{ex}} - \phi^\dagger_{\mathrm{ex}}\varpi_{\mathrm{ex}}\phi_{\mathrm{ex}} + \frac{d}{dt}(\phi^\dagger_{\mathrm{ex}}\dot{m}_{\mathrm{ex}}\phi_{\mathrm{ex}})\right] \tag{IV.5.5}$$

  where m_{ex} and ϖ_{ex} are hermitian matrices acting on $\mathbb{C}^{\mathcal{S}}$ with the form

  $$(m_{\mathrm{ex}})_{\mathbf{M}_1\mathbf{M}_2} = \delta_{\mathbf{M}_1\mathbf{M}_2} - \lambda\Phi_{\mathbf{M}_1-\mathbf{M}_2} \tag{IV.5.6a}$$
  $$(\varpi_{\mathrm{ex}})_{\mathbf{M}_1\mathbf{M}_2} = \delta_{\mathbf{M}_1\mathbf{M}_2}\mathbf{k}^2_{\mathbf{M}_1} - \lambda(k^2_{\mathbf{M}_1\mathbf{M}_2}\Phi_{\mathbf{M}_1-\mathbf{M}_2} + \ddot{\Phi}_{\mathbf{M}_1-\mathbf{M}_2}). \tag{IV.5.6b}$$

  Unfortunately, the components of ϕ_{ex} represent twice as many degrees of freedom as are integrated over in (IV.5.4). This means that a path integral over all the components would have to include the factor

  $$\prod_{\mathbf{M}\in\mathcal{S}/2}\delta\left(\phi_{-\mathbf{M}} - \phi^*_{\mathbf{M}}\right). \tag{IV.5.7}$$

  This leads one to consider

- a complex vector ϕ_+ with components $\{\phi_{\mathbf{M}}|\mathbf{M} \in \mathcal{S}/2\}$. (We similarly define this space as $\mathbb{C}^{\mathcal{S}/2}$.) This would completely specify the unique modes of ϕ, but (IV.5.2) is not conveniently expressed in terms of ϕ_+. To see this, consider the velocity term

$$\begin{aligned}\frac{1}{2}&\sum_{\mathbf{M}_1,\mathbf{M}_2\in\mathcal{S}} \dot{\phi}^*_{\mathbf{M}_1}(\delta_{\mathbf{M}_1\mathbf{M}_2} - \lambda\Phi_{\mathbf{M}_1-\mathbf{M}_2})\dot{\phi}_{\mathbf{M}_2} \\ &= \frac{1}{2}\sum_{\mathbf{M}_1,\mathbf{M}_2\in\mathcal{S}/2} \big[\dot{\phi}^*_{\mathbf{M}_1}(\delta_{\mathbf{M}_1\mathbf{M}_2} - \lambda\Phi_{\mathbf{M}_1-\mathbf{M}_2})\dot{\phi}_{\mathbf{M}_2} + \dot{\phi}_{\mathbf{M}_1}(\delta_{\mathbf{M}_1\mathbf{M}_2} - \lambda\Phi_{\mathbf{M}_2-\mathbf{M}_1})\dot{\phi}^*_{\mathbf{M}_2} \\ &\qquad - \dot{\phi}^*_{\mathbf{M}_1}\lambda\Phi_{\mathbf{M}_1+\mathbf{M}_2}\dot{\phi}^*_{\mathbf{M}_2} - \dot{\phi}_{\mathbf{M}_1}\lambda\Phi^*_{\mathbf{M}_1+\mathbf{M}_2}\dot{\phi}_{\mathbf{M}_2}\big] \end{aligned} \tag{IV.5.8}$$

Although the first two terms can be written as

$$\sum_{\mathbf{M}_1,\mathbf{M}_2\in\mathcal{S}/2} \dot{\phi}^*_{\mathbf{M}_1}(\delta_{\mathbf{M}_1\mathbf{M}_2} - \lambda\Phi_{\mathbf{M}_1-\mathbf{M}_2})\dot{\phi}_{\mathbf{M}_2} = \phi_+^\dagger m_+ \phi_+, \tag{IV.5.9}$$

the last two give

$$\mathrm{Re}\left[\sum_{\mathbf{M}_1,\mathbf{M}_2\in\mathcal{S}/2} \dot{\phi}_{\mathbf{M}_1}(\delta_{\mathbf{M}_1\mathbf{M}_2} - \lambda\Phi^*_{\mathbf{M}_1+\mathbf{M}_2})\dot{\phi}_{\mathbf{M}_2},\right] \tag{IV.5.10}$$

which cannot be written in terms of the complex vector ϕ_+ and its adjoint $\phi_+^\dagger$ without using the transpose ϕ_+^{tr} or the complex conjugate ϕ_+^*. If $D = 1$, this is not a problem, since $\mathbf{M}_1, \mathbf{M}_2 \in \mathcal{S}/2$ implies $\mathbf{M}_1 + \mathbf{M}_2 \notin \mathcal{L}$ and hence $\Phi_{\mathbf{M}_1+\mathbf{M}_2} = 0$. However, it is possible in $D > 1$ to have $\mathbf{M}_1 + \mathbf{M}_2 \in \mathcal{L}$ even when $\mathbf{M}_1, \mathbf{M}_2 \in \mathcal{S}/2$, as illustrated in Fig. IV.1. This leaves

- a real vector ϕ with components $\{\phi^{\mathrm{R}}_{\mathbf{M}}, \phi^{\mathrm{I}}_{\mathbf{M}}|\mathbf{M} \in \mathcal{S}/2\}$ (which lies in the space we define as $\mathbb{R}^{\mathcal{S}/2} \otimes \mathbb{R}^{\mathcal{S}/2} = \mathbb{R}^{\mathcal{S}/2\oplus\mathcal{S}/2}$). This method is basically fool-proof. Given a complex vector $v_{\mathrm{ex}} \in \mathbb{C}^{\mathcal{S}}$ and a hermitian matrix M_{ex} [which also obeys $(M_{\mathrm{ex}})_{-\mathbf{M}_1,-\mathbf{M}_2} = (M_{\mathrm{ex}})^*_{\mathbf{M}_1,\mathbf{M}_2}$], we have

$$v^\dagger_{\mathrm{ex}} M_{\mathrm{ex}} v_{\mathrm{ex}} = (v^{\mathrm{R}}_{\mathrm{ex}} - iv^{\mathrm{I}}_{\mathrm{ex}})^{\mathrm{tr}}(M^{\mathrm{R}}_{\mathrm{ex}} + iM^{\mathrm{I}}_{\mathrm{ex}})(v^{\mathrm{R}}_{\mathrm{ex}} + iv^{\mathrm{I}}_{\mathrm{ex}}) = \begin{pmatrix} v^{\mathrm{R}}_{\mathrm{ex}} \\ v^{\mathrm{I}}_{\mathrm{ex}}\end{pmatrix}^{\mathrm{tr}} \begin{pmatrix} M^{\mathrm{R}}_{\mathrm{ex}} & -M^{\mathrm{I}}_{\mathrm{ex}} \\ M^{\mathrm{I}}_{\mathrm{ex}} & M^{\mathrm{R}}_{\mathrm{ex}}\end{pmatrix} \begin{pmatrix} v^{\mathrm{R}}_{\mathrm{ex}} \\ v^{\mathrm{I}}_{\mathrm{ex}};\end{pmatrix} \tag{IV.5.11}$$

If we define complex vectors $v_\pm = \{(v_{\mathrm{ex}})_{\pm\mathbf{M}}|\mathbf{M} \in \mathcal{S}/2\}$ in, and matrices[2] $M_\pm = \{(M_{\mathrm{ex}})_{\mathbf{M}_1,\pm\mathbf{M}_2}|\mathbf{M}_1,\mathbf{M}_2 \in \mathcal{S}/2\}$ acting on, $\mathbb{C}^{\mathcal{S}/2}$, we can write the quantities involving $\mathbb{C}^{\mathcal{S}}$ in terms of them as $v_{\mathrm{ex}} = \left(\begin{smallmatrix} v_+ \\ v_- \end{smallmatrix}\right)$ and $M_{\mathrm{ex}} = \left(\begin{smallmatrix} M_+ & M_- \\ M_-^\dagger & M_+^* \end{smallmatrix}\right)$ or

$$M^{\mathrm{R}}_{\mathrm{ex}} = \begin{pmatrix} M^{\mathrm{R}}_+ & M^{\mathrm{R}}_- \\ M^{\mathrm{R}\,\mathrm{tr}}_- & M^{\mathrm{R}}_+ \end{pmatrix} = \begin{pmatrix} M^{\mathrm{R}}_+ & M^{\mathrm{R}}_- \\ M^{\mathrm{R}}_- & M^{\mathrm{R}}_+ \end{pmatrix} \tag{IV.5.12a}$$

$$M^{\mathrm{I}}_{\mathrm{ex}} = \begin{pmatrix} M^{\mathrm{I}}_+ & M^{\mathrm{I}}_- \\ -M^{\mathrm{I}\,\mathrm{tr}}_- & -M^{\mathrm{I}}_+ \end{pmatrix} = \begin{pmatrix} M^{\mathrm{I}}_+ & M^{\mathrm{I}}_- \\ -M^{\mathrm{I}}_- & -M^{\mathrm{I}}_+ \end{pmatrix}. \tag{IV.5.12b}$$

[2] It is easy to work out that while M_+ is hermitian by the hermiticity of M_{ex}, M_- is symmetric because $(M_{\mathrm{ex}})_{\mathbf{M}_1,-\mathbf{M}_2} = (M_{\mathrm{ex}})^*_{-\mathbf{M}_1,\mathbf{M}_2} = (M_{\mathrm{ex}})_{\mathbf{M}_2,-\mathbf{M}_1}$.

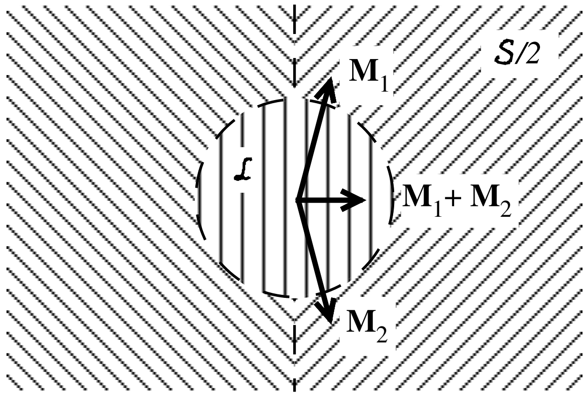


Figure IV.1: The addition of momenta $\mathbf{M}_1, \mathbf{M}_2 \in \mathcal{S}/2$ can produce $\mathbf{M}_1 + \mathbf{M}_2 \in \mathcal{L}$. The long-wavelength (low-momentum) region $\mathcal{L}$ is shaded vertically. The short-wavelength (high-momentum) region $\mathcal{S}$ is shaded diagonally in one direction or the other. The right half of $\mathcal{S}$, shaded diagonally up and to the right, is $\mathcal{S}/2$. In $D > 1$, we see that it is possible to add two "large" momenta on the right side of the origin ($\mathbf{M}_1, \mathbf{M}_2 \in \mathcal{S}/2$) to get a "small" momentum ($\mathbf{M}_1 + \mathbf{M}_2 \in \mathcal{L}$).

The reality condition $v_{-\mathbf{M}} = v^*_{\mathbf{M}}$ becomes $v_- = v^*_+$, so that

$$\begin{pmatrix} v^{\mathrm{R}}_{\mathrm{ex}} \\ v^{\mathrm{I}}_{\mathrm{ex}} \end{pmatrix} = \begin{pmatrix} v^{\mathrm{R}}_+ \\ v^{\mathrm{R}}_- \\ v^{\mathrm{I}}_+ \\ v^{\mathrm{I}}_- \end{pmatrix} = \begin{pmatrix} v^{\mathrm{R}}_+ \\ v^{\mathrm{R}}_+ \\ v^{\mathrm{I}}_+ \\ -v^{\mathrm{I}}_+ \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} v^{\mathrm{R}}_+ \\ v^{\mathrm{I}}_+ \end{pmatrix}. \tag{IV.5.13}$$

We can combine this with (IV.5.11) and (IV.5.12) to give

$$\begin{aligned} {v_{\mathrm{ex}}}^{\mathrm{tr}} M_{\mathrm{ex}} v_{\mathrm{ex}} =& \begin{pmatrix} v^{\mathrm{R}}_+ \\ v^{\mathrm{I}}_+ \end{pmatrix}^{\mathrm{tr}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} M^{\mathrm{R}}_+ & M^{\mathrm{R}}_- & -M^{\mathrm{I}}_+ & -M^{\mathrm{I}}_- \\ M^{\mathrm{R}}_- & M^{\mathrm{R}}_+ & M^{\mathrm{I}}_- & M^{\mathrm{I}}_+ \\ M^{\mathrm{I}}_+ & M^{\mathrm{I}}_- & M^{\mathrm{R}}_+ & M^{\mathrm{R}}_- \\ -M^{\mathrm{I}}_- & -M^{\mathrm{I}}_+ & M^{\mathrm{R}}_- & M^{\mathrm{R}}_+ \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} v^{\mathrm{R}}_+ \\ v^{\mathrm{I}}_+ \end{pmatrix} \\ =& 2 \begin{pmatrix} v^{\mathrm{R}}_+ \\ v^{\mathrm{I}}_+ \end{pmatrix}^{\mathrm{tr}} \begin{pmatrix} M^{\mathrm{R}}_+ + M^{\mathrm{R}}_- & -M^{\mathrm{I}}_+ + M^{\mathrm{I}}_- \\ M^{\mathrm{I}}_+ + M^{\mathrm{I}}_- & M^{\mathrm{R}}_+ - M^{\mathrm{R}}_- \end{pmatrix} \begin{pmatrix} v^{\mathrm{R}}_+ \\ v^{\mathrm{I}}_+ \end{pmatrix}. \end{aligned} \tag{IV.5.14}$$

This allows us to rewrite (IV.5.5) as

$$L_{0+\phi\phi} = \frac{1}{2}\left[\dot{\phi}^{\mathrm{tr}} m \dot{\phi} - \phi^{\mathrm{tr}} \varpi \phi + \frac{d}{dt}(\phi^{\mathrm{tr}} \dot{m} \phi)\right], \tag{IV.5.15}$$

where

$$\phi = \sqrt{2}\begin{pmatrix}\phi_+^{\mathrm{R}} \\ \phi_+^{\mathrm{I}}\end{pmatrix} = \sqrt{2}\begin{pmatrix}\{\phi_{\mathbf{M}}^{\mathrm{R}}\} \\ \{\phi_{\mathbf{M}}^{\mathrm{I}}\}\end{pmatrix} \tag{IV.5.16}$$

$$m = \begin{pmatrix}\{\delta_{\mathbf{M}_1\mathbf{M}_2} - \lambda(\Phi^{\mathrm{R}}_{\mathbf{M}_1-\mathbf{M}_2} + \Phi^{\mathrm{R}}_{\mathbf{M}_1+\mathbf{M}_2})\} & \{-\lambda(-\Phi^{\mathrm{I}}_{\mathbf{M}_1-\mathbf{M}_2} + \Phi^{\mathrm{I}}_{\mathbf{M}_1+\mathbf{M}_2})\} \\ \{-\lambda(\Phi^{\mathrm{I}}_{\mathbf{M}_1-\mathbf{M}_2} + \Phi^{\mathrm{I}}_{\mathbf{M}_1+\mathbf{M}_2})\} & \{\delta_{\mathbf{M}_1\mathbf{M}_2} - \lambda(\Phi^{\mathrm{R}}_{\mathbf{M}_1-\mathbf{M}_2} - \Phi^{\mathrm{R}}_{\mathbf{M}_1+\mathbf{M}_2})\}\end{pmatrix} \tag{IV.5.17}$$

$$\varpi = \begin{pmatrix}\{\varpi^{\mathsf{UL}}_{\mathbf{M}_1\mathbf{M}_2}\} & \{\varpi^{\mathsf{UR}}_{\mathbf{M}_1\mathbf{M}_2}\} \\ \{\varpi^{\mathsf{LL}}_{\mathbf{M}_1\mathbf{M}_2}\} & \{\varpi^{\mathsf{LR}}_{\mathbf{M}_1\mathbf{M}_2}\}\end{pmatrix} \tag{IV.5.18}$$

and

$$\varpi^{\mathsf{UL}}_{\mathbf{M}_1\mathbf{M}_2} = \delta_{\mathbf{M}_1\mathbf{M}_2}\mathbf{k}^2_{\mathbf{M}_1} - \lambda(\Phi^{\mathrm{R}}_{\mathbf{M}_1-\mathbf{M}_2}k^2_{\mathbf{M}_1\mathbf{M}_2} + \Phi^{\mathrm{R}}_{\mathbf{M}_1+\mathbf{M}_2}k^2_{\mathbf{M}_1,-\mathbf{M}_2} + \ddot{\Phi}^{\mathrm{R}}_{\mathbf{M}_1-\mathbf{M}_2} + \ddot{\Phi}^{\mathrm{R}}_{\mathbf{M}_1+\mathbf{M}_2}) \tag{IV.5.18a}$$

$$\varpi^{\mathsf{UR}}_{\mathbf{M}_1\mathbf{M}_2} = -\lambda(-\Phi^{\mathrm{I}}_{\mathbf{M}_1-\mathbf{M}_2}k^2_{\mathbf{M}_1\mathbf{M}_2} + \Phi^{\mathrm{I}}_{\mathbf{M}_1+\mathbf{M}_2}k^2_{\mathbf{M}_1,-\mathbf{M}_2} - \ddot{\Phi}^{\mathrm{I}}_{\mathbf{M}_1-\mathbf{M}_2} + \ddot{\Phi}^{\mathrm{I}}_{\mathbf{M}_1+\mathbf{M}_2}) \tag{IV.5.18b}$$

$$\varpi^{\mathsf{LL}}_{\mathbf{M}_1\mathbf{M}_2} = -\lambda(\Phi^{\mathrm{I}}_{\mathbf{M}_1-\mathbf{M}_2}k^2_{\mathbf{M}_1\mathbf{M}_2} + \Phi^{\mathrm{I}}_{\mathbf{M}_1+\mathbf{M}_2}k^2_{\mathbf{M}_1,-\mathbf{M}_2} + \ddot{\Phi}^{\mathrm{I}}_{\mathbf{M}_1-\mathbf{M}_2} + \ddot{\Phi}^{\mathrm{I}}_{\mathbf{M}_1+\mathbf{M}_2}) \tag{IV.5.18c}$$

$$\varpi^{\mathsf{LR}}_{\mathbf{M}_1\mathbf{M}_2} = \delta_{\mathbf{M}_1\mathbf{M}_2}\mathbf{k}^2_{\mathbf{M}_1} - \lambda(\Phi^{\mathrm{R}}_{\mathbf{M}_1-\mathbf{M}_2}k^2_{\mathbf{M}_1\mathbf{M}_2} - \Phi^{\mathrm{R}}_{\mathbf{M}_1+\mathbf{M}_2}k^2_{\mathbf{M}_1,-\mathbf{M}_2} + \ddot{\Phi}^{\mathrm{R}}_{\mathbf{M}_1-\mathbf{M}_2} - \ddot{\Phi}^{\mathrm{R}}_{\mathbf{M}_1+\mathbf{M}_2}); \tag{IV.5.18d}$$

ϕ is a vector in, and m and ϖ are real symmetric matrices acting on, $\mathbb{R}^{\mathcal{S}/2\oplus\mathcal{S}/2}$.

In light of (IV.5.15), the third option is the most useful one. If we choose the normalization of (IV.5.4) so that

$$\mathcal{D}\phi = \prod_{\mathbf{M}\in\mathcal{S}/2} 2\mathcal{D}\phi^{\mathrm{R}}_{\mathbf{M}}\mathcal{D}\phi^{\mathrm{I}}_{\mathbf{M}}, \tag{IV.5.19}$$

The measure $\mathcal{D}\phi$ is just the product of the measures corresponding to all the components of ϕ.

IV.5.2 The propagator

We now have a workable vector expression for the path integral (IV.5.1):

$$\begin{aligned} e^{iW_{0+\phi\phi}[\Phi_1,\Phi_2]} &= \int \mathcal{D}\phi_1\mathcal{D}\phi_2\rho_\phi(\phi_1',\phi_2')\delta(\phi_2''-\phi_1'')e^{i(S_{0+\phi\phi}[\phi_1,\Phi_1]-S_{0+\phi\phi}[\phi_2,\Phi_2])} \\ &= \int d\phi_1' d\phi_2' d\phi'' \rho_\phi(\phi_1',\phi_2')\mathcal{K}_{0+\phi\phi}(\phi''|\phi_1';\Phi_1]\mathcal{K}^*_{0+\phi\phi}(\phi''|\phi_2';\Phi_2] \end{aligned} \tag{IV.5.20}$$

where

$$\mathcal{K}_{0+\phi\phi}(\phi''|\phi';\Phi] = \int\limits_{\phi''\phi'} \mathcal{D}\phi e^{iS_{0+\phi\phi}[\phi,\Phi]} \tag{IV.5.21}$$

is the propagator for the quadratic action. It is useful to write

$$\mathcal{K}_{0+\phi\phi}(\phi''|\phi';\Phi] = e^{\frac{i}{2}(\phi''^{\text{tr}}\dot{m}''\phi''-\phi'^{\text{tr}}\dot{m}'\phi')} K\left(\phi''\frac{T}{2}\,\middle|\,\phi'\frac{T}{2}\right) \tag{IV.5.22}$$

where

$$K(\phi_b t_b|\phi_a t_a) = \int\limits_{\phi_b\phi_a} e^{\frac{i}{2}\int_{t_a}^{t_b} dt[\dot{\phi}(t)^{\text{tr}}m(t)\dot{\phi}(t)-\phi(t)^{\text{tr}}\varpi(t)\phi(t)]} \tag{IV.5.23}$$

is the propagator for a simple harmonic oscillator with time-dependent matrices $m(t)$ and $\varpi(t)$ in place of m and $m\omega^2$. [The dependence on Φ is now implicit in the time dependence of $m(t)$ and $\varpi(t)$, given by (IV.5.17–IV.5.18).]

This propagator is found in Appendix IV.B to be

$$K(\phi_b t_b|\phi_a t_a) = \frac{1}{\sqrt{\det\left(2\pi i\mathfrak{C}(t_b|t_a)\right)}} \exp\left[\frac{i}{2}\begin{pmatrix}\phi_b\\ \phi_a\end{pmatrix}^{\text{tr}}\begin{pmatrix}\mathfrak{C}^{-1}(t_b|t_a)\mathfrak{B}(t_b|t_a) & -\mathfrak{C}^{-1}(t_b|t_a)\\ -\mathfrak{C}^{-1}(t_b|t_a)^{\text{tr}} & \mathfrak{A}(t_b|t_a)\mathfrak{C}^{-1}(t_b|t_a)\end{pmatrix}\begin{pmatrix}\phi_b\\ \phi_a\end{pmatrix}\right], \tag{IV.B.8$'$}$$

where

$$\mathfrak{B}(t_b|t_a) = \sum_{n=0}^{\infty}\left(\prod_{k=1}^{n}\int_{t_a}^{\tilde{t}_{k-1}} dt_k \int_{t_a}^{t_k} d\tilde{t}_k\right)\prod_{k=n}^{1}\left[-m^{-1}(\tilde{t}_k)\varpi(t_k)\right] \tag{IV.B.19a}$$

$$\mathfrak{C}(t_b|t_a) = -\frac{d\mathfrak{B}(t_b|t_a)}{dt_b}\varpi^{-1}(t_b) \tag{IV.B.19b}$$

$$\mathfrak{A}(t_b|t_a) = -m(t_a)\frac{d\mathfrak{C}(t_b|t_a)}{dt_a}. \tag{IV.B.19c}$$

These exact expressions are expanded to first order in λ in Sec. IV.C.1 of Appendix IV.C, using the values of $m(t)$ and $\varpi(t)$ given by (IV.5.17) and (IV.5.18), respectively.

Given the expression (IV.B.8$'$) for the time-dependent propagator, (IV.5.22) becomes

$$\mathcal{K}_{0+\phi\phi}(\phi''|\phi';\Phi] = \frac{1}{\sqrt{\det(2\pi i\mathfrak{C}[\Phi])}} \exp\left[\frac{i}{2}\begin{pmatrix}\phi''\\ \phi'\end{pmatrix}^{\text{tr}}\begin{pmatrix}B[\Phi] & -C[\Phi]\\ -C[\Phi]^{\text{tr}} & A[\Phi]\end{pmatrix}\begin{pmatrix}\phi''\\ \phi'\end{pmatrix}\right], \tag{IV.5.24}$$

where

$$A[\Phi] = \mathfrak{A}\left(\frac{T}{2}\,\middle|-\frac{T}{2}\right)\mathfrak{C}^{-1}\left(\frac{T}{2}\,\middle|-\frac{T}{2}\right) - \dot{m}\left(-\frac{T}{2}\right) \tag{IV.5.25a}$$

$$B[\Phi] = \mathfrak{C}^{-1}\left(\frac{T}{2}\,\middle|-\frac{T}{2}\right)\mathfrak{B}\left(\frac{T}{2}\,\middle|-\frac{T}{2}\right) + \dot{m}\left(\frac{T}{2}\right) \tag{IV.5.25b}$$

$$C[\Phi] = \mathfrak{C}^{-1}\left(\frac{T}{2}\,\middle|-\frac{T}{2}\right). \tag{IV.5.25c}$$

This means that (IV.5.20) becomes

$$
\begin{aligned}
e^{iW_{0+\phi\phi}[\Phi_1,\Phi_2]} = &\int \frac{d\phi_1' d\phi_2' d\phi'' \rho_\phi(\phi_1',\phi_2')}{\sqrt{\det(2\pi\mathfrak{C}[\Phi_1])\det(2\pi\mathfrak{C}[\Phi_2])}} \\
&\times \exp\left[\frac{i}{2}\begin{pmatrix}\phi''\\ \phi_1'\\ \phi_2'\end{pmatrix}^{\mathrm{tr}}\begin{pmatrix}B[\Phi_1]-B[\Phi_2] & -C[\Phi_1] & C[\Phi_2]\\ -C[\Phi_1]^{\mathrm{tr}} & A[\Phi_1] & 0\\ C[\Phi_2]^{\mathrm{tr}} & 0 & -A[\Phi_2]\end{pmatrix}\begin{pmatrix}\phi''\\ \phi_1'\\ \phi_2'\end{pmatrix}\right]
\end{aligned}
\tag{IV.5.26}
$$

IV.5.3 The initial state

For the initial state ρ_ϕ of the SWMs I choose a thermal state with temperature $1/k_{\mathrm{B}}\beta$. The density matrix for this is given as an operator by $\widehat{\rho} \propto e^{-\beta\widehat{H}}$. Using the full hamiltonian corresponding to the action (IV.4.9) would couple the short- and long-wavelength modes, preventing the separation (IV.2.3) of the initial state. So instead I use the zero-order non-interacting action S_0, which gives the thermal density matrix for a simple harmonic oscillator of frequency $\Omega_0 = \left(\begin{smallmatrix}\mathrm{diag}\{k_{\mathbf{M}}\} & 0\\ 0 & \mathrm{diag}\{k_{\mathbf{M}}\}\end{smallmatrix}\right)$ and unit mass:

$$
\rho_\phi(\phi_1',\phi_2') \propto \exp\left[-\frac{1}{2}\begin{pmatrix}\phi_1'\\ \phi_2'\end{pmatrix}^{\mathrm{tr}}\begin{pmatrix}\frac{\Omega_0}{\tanh\Omega_0\beta} & -\frac{\Omega_0}{\sinh\Omega_0\beta}\\ -\frac{\Omega_0}{\sinh\Omega_0\beta} & \frac{\Omega_0}{\tanh\Omega_0\beta}\end{pmatrix}\begin{pmatrix}\phi_1'\\ \phi_2'\end{pmatrix}\right]
\tag{IV.5.27}
$$

Equation (IV.5.24) is simplified if we express it in terms of $\overline{\phi'} = \frac{\phi_1'+\phi_2'}{2}$ and $\Delta\phi' = \phi_1' - \phi_2'$ using

$$
\begin{pmatrix}\phi_1'\\ \phi_2'\end{pmatrix} = \begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\begin{pmatrix}\overline{\phi'}\\ \Delta\phi'/2\end{pmatrix}.
\tag{IV.5.28}
$$

Both $\frac{\Omega_0}{\tanh\Omega_0\beta} - \frac{\Omega_0}{\sinh\Omega_0\beta}$ and $\frac{\Omega_0}{\tanh\Omega_0\beta} + \frac{\Omega_0}{\sinh\Omega_0\beta}$ can be expressed in terms of

$$
\mathcal{V}(\Omega_0) = \frac{2}{\Omega_0}\frac{\cosh\Omega_0\beta - 1}{\sinh\Omega_0\beta} = \frac{2}{\Omega_0}\frac{\sinh\Omega_0\beta}{\cosh\Omega_0\beta + 1} = \frac{2}{\Omega_0}\sqrt{\frac{\cosh\Omega_0\beta - 1}{\cosh\Omega_0\beta + 1}} = \frac{2}{\Omega_0}\tanh\frac{\Omega_0\beta}{2}
\tag{IV.5.29}
$$

to give

$$
\rho_\phi(\phi_1',\phi_2') \propto \exp\left[-\frac{1}{2}\begin{pmatrix}\overline{\phi'}\\ \Delta\phi'/2\end{pmatrix}^{\mathrm{tr}}\begin{pmatrix}\Omega_0^2\mathcal{V}(\Omega_0) & 0\\ 0 & \mathcal{V}^{-1}(\Omega_0)\end{pmatrix}\begin{pmatrix}\overline{\phi'}\\ \Delta\phi'/2\end{pmatrix}\right];
\tag{IV.5.30}
$$

if we also define

$$
\mathcal{A}_\pm = A[\Phi_1] \pm A[\Phi_2] \tag{IV.5.31a}
$$
$$
\mathcal{B}_\pm = B[\Phi_1] \pm B[\Phi_2] \tag{IV.5.31b}
$$
$$
\mathcal{C}_\pm = C[\Phi_1] \pm C[\Phi_2], \tag{IV.5.31c}
$$

(IV.5.26) becomes

$$
\begin{aligned}
e^{iW_{0+\phi\phi}[\Phi_1,\Phi_2]} &\propto \int \frac{d\overline{\phi'}\,d\Delta\phi'\,d\phi''}{\sqrt{\det(2\pi\mathfrak{C}[\Phi_1])\det(2\pi\mathfrak{C}[\Phi_2])}} \exp\left[-\frac{1}{2}\begin{pmatrix}\phi''\\ \overline{\phi'}\\ \Delta\phi'/2\end{pmatrix}^{\mathrm{tr}} \mathcal{M} \begin{pmatrix}\phi''\\ \overline{\phi'}\\ \Delta\phi'/2\end{pmatrix}\right] \\
&= \left\{\det(2\pi\mathfrak{C}[\Phi_1])\det(2\pi\mathfrak{C}[\Phi_2])\det\left(\frac{\mathcal{M}}{2\pi}\right)\right\}^{-1/2}
\end{aligned}
\tag{IV.5.32}
$$

where

$$
\mathcal{M} = \begin{pmatrix} -i\mathcal{B}_- & i\mathcal{C}_- & i\mathcal{C}_+ \\ i{\mathcal{C}_-}^{\mathrm{tr}} & \Omega_0^2\mathcal{V}(\Omega_0) - i\mathcal{A}_- & -i\mathcal{A}_+ \\ i{\mathcal{C}_+}^{\mathrm{tr}} & -i\mathcal{A}_+ & 4\mathcal{V}^{-1}(\Omega_0) - i\mathcal{A}_- \end{pmatrix} \tag{IV.5.33}
$$

Since the matrix $A[\Phi]$ can be expanded (see Sec. IV.C.2 of Appendix IV.C) as $A[\Phi] = A_0 + \lambda A_1[\Phi] + \mathcal{O}(\lambda^2)$, where $A_1[\Phi]$ is a linear function of its argument, we have $\mathcal{A}_+ = 2A_0 + \mathcal{O}(\lambda)$ and $\mathcal{A}_- = \lambda A_1[\Delta\Phi] + \mathcal{O}(\lambda^2)$, with similar expressions holding for $\mathcal{B}_\pm$ and $\mathcal{C}_\pm$. This means the sub-matrices of $\mathcal{M}$ are of the following order:

$$
\mathcal{M} = \begin{pmatrix} \mathcal{O}(\lambda) & \mathcal{O}(\lambda) & \mathcal{O}(1) \\ \mathcal{O}(\lambda) & \Omega_0^2\mathcal{V}(\Omega_0) + \mathcal{O}(\lambda) & \mathcal{O}(1) \\ \mathcal{O}(1) & \mathcal{O}(1) & 4\mathcal{V}^{-1}(\Omega_0) + \mathcal{O}(\lambda) \end{pmatrix}. \tag{IV.5.34}
$$

Given the relation

$$
4\mathcal{V}^{-1}(\Omega_0) = 2\Omega_0\sqrt{\frac{\cosh\Omega_0\beta + 1}{\cosh\Omega_0\beta - 1}} \geq 2\Omega_0 \geq 2\Omega_0\sqrt{\frac{\cosh\Omega_0\beta - 1}{\cosh\Omega_0\beta + 1}} = \Omega_0^2\mathcal{V}(\Omega_0) \tag{IV.5.35}
$$

we see that $\alpha = 4\mathcal{V}^{-1}(\Omega_0) - i\mathcal{A}_-$ is the largest of the sub-matrices on the diagonal, and is no smaller than $\mathcal{O}(1)$. Thus we partially diagonalize $\mathcal{M}$ about it to get

$$
\begin{aligned}
\widetilde{\mathcal{M}} &= \begin{pmatrix} 1 & 0 & -i\mathcal{C}_+\alpha^{-1} \\ 0 & 1 & i\mathcal{A}_+\alpha^{-1} \\ 0 & 0 & 1 \end{pmatrix} \mathcal{M} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -i\alpha^{-1}{\mathcal{C}_+}^{\mathrm{tr}} & i\alpha^{-1}\mathcal{A}_+ & 1 \end{pmatrix} \\
&= \begin{pmatrix} \mathcal{C}_+\alpha^{-1}{\mathcal{C}_+}^{\mathrm{tr}} - i\mathcal{B}_- & -\mathcal{C}_+\alpha^{-1}\mathcal{A}_+ + i\mathcal{C}_- & 0 \\ -\mathcal{A}_+\alpha^{-1}{\mathcal{C}_+}^{\mathrm{tr}} + i{\mathcal{C}_-}^{\mathrm{tr}} & \mathcal{A}_+\alpha^{-1}\mathcal{A}_+ + \Omega_0^2\mathcal{V}(\Omega_0) - i\mathcal{A}_- & 0 \\ 0 & 0 & \alpha \end{pmatrix}
\end{aligned}
\tag{IV.5.36}
$$

IV.5.4 Controlling the breakdown of perturbation theory

Before proceeding further, we need to consider more carefully the perturbation theory approach. If we try to expand the influence functional e^{iW} defined by (IV.2.4b) in powers of λ, we note that as the zero-order term in $S_E[\phi,\Phi]$ is just $S_0[\phi]$ [*i.e.*, the "system" and "environment" are decoupled to zeroth order; *cf.* (IV.4.9)],

$$
\left(e^{iW[\Phi_1,\Phi_2]}\right)_0 = \int \mathcal{D}\phi_1\mathcal{D}\phi_2\rho_\phi(\phi_1',\phi_2')\delta(\phi_2''-\phi_1'')e^{i(S_0[\phi_1]-S_0[\phi_2])} = \mathrm{Tr}\left[e^{-iT\widehat{H}}\rho_\phi e^{iT\widehat{H}}\right] = 1. \tag{IV.5.37}
$$

Perturbatively, then, we would conclude $e^{iW} = 1 + \mathcal{O}(\lambda)$. The problem is that for the influence phase to be effective at producing decoherence, we need $e^{iW[\Phi_1\Phi_2]} \ll 1$ for sufficiently different coarse grainings. This can only be possible if the perturbative analysis breaks down somehow. I will focus my attention on a scenario where that breakdown is manageable. If the temperature β^{-1} is high enough, there will be some modes in $\mathcal{S}$ for which $\mathcal{V}(k) = \frac{2}{k}\frac{\sinh k\beta}{\cosh k\beta+1} \to \beta$, and the $\mathcal{O}(\beta)$ terms like $\mathcal{C}_+\alpha^{-1}\mathcal{C}_+{}^{\mathrm{tr}}$ may become smaller than $\mathcal{O}(\lambda)$ terms like $\mathcal{B}_-$. At that point, if the $\mathcal{O}(\lambda)$ correction to e^{iW} is also $\mathcal{O}(\beta^{-1})$, it can cause perturbation theory to break down. We keep a handle on this breakdown by neglecting $\mathcal{O}(\lambda)$ terms only when they are not compared to potentially $\mathcal{O}(\beta)$ terms.

IV.5.5 Evaluation of the influence phase

Using the approximation of Sec. IV.5.4, we have

$$\begin{aligned}\widetilde{\mathcal{M}} &= \begin{pmatrix} \mathcal{C}_+\frac{\mathcal{V}(\Omega_0)}{4}\mathcal{C}_+{}^{\mathrm{tr}} - i\mathcal{B}_- & -\mathcal{C}_+\frac{\mathcal{V}(\Omega_0)}{4}\mathcal{A}_+ + i\mathcal{C}_- \\ -\mathcal{A}_+\frac{\mathcal{V}(\Omega_0)}{4}\mathcal{C}_+{}^{\mathrm{tr}} + i\mathcal{C}_-{}^{\mathrm{tr}} & \mathcal{A}_+\frac{\mathcal{V}(\Omega_0)}{4}\mathcal{A}_+ + \Omega_0^2\mathcal{V}(\Omega_0) - i\mathcal{A}_- \end{pmatrix} \oplus 4\mathcal{V}^{-1}(\Omega_0) \\ &= \begin{pmatrix} \frac{\Omega_0^2\mathcal{V}(\Omega_0)}{\sin^2\Omega_0 T} - i\lambda B_1[\Delta\Phi] & -\frac{\Omega_0^2\mathcal{V}(\Omega_0)\cos\Omega_0 T}{\sin^2\Omega_0 T} + i\lambda C_1[\Delta\Phi] \\ -\frac{\Omega_0^2\mathcal{V}(\Omega_0)\cos\Omega_0 T}{\sin^2\Omega_0 T} + i\lambda C_1[\Delta\Phi]^{\mathrm{tr}} & \frac{\Omega_0^2\mathcal{V}(\Omega_0)}{\sin^2\Omega_0 T} - i\lambda A_1[\Delta\Phi] \end{pmatrix} \oplus 4\mathcal{V}^{-1}(\Omega_0).\end{aligned} \tag{IV.5.38}$$

Noting that the matrices used to perform the diagonalization in (IV.5.36) have unit determinant, we have

$$\det\mathcal{M} = \det\widetilde{\mathcal{M}} = \det\left(4\mathcal{V}^{-1}(\Omega_0)\right)\det(\aleph_0 - i\lambda\aleph_1[\Delta\Phi]) \propto \det\left(1 - i\lambda\aleph_0^{-1/2}\aleph_1[\Delta\Phi]\aleph_0^{-1/2}\right) \tag{IV.5.39}$$

where

$$\aleph_0 = \frac{\Omega_0^2\mathcal{V}(\Omega_0)}{\sin^2\Omega_0 T}\begin{pmatrix} 1 & -\cos\Omega_0 T \\ -\cos\Omega_0 T & 1 \end{pmatrix} \quad \text{and} \quad \aleph_1[\Delta\Phi] = \begin{pmatrix} B_1[\Delta\Phi] & -C_1[\Delta\Phi] \\ -C_1[\Delta\Phi]^{\mathrm{tr}} & A_1[\Delta\Phi] \end{pmatrix} \tag{IV.5.40}$$

Now, $e^{i\,\mathrm{Re}\,W}$ is simply a phase multiplying the decoherence functional (IV.2.4); the part which can actually make the off-diagonal components of $D[\Phi_1,\Phi_2]$ small is $e^{-\mathrm{Im}\,W} = \left|e^{iW}\right|$. Noting that the factors of $\det\mathfrak{C}$ in (IV.5.32) give, to lowest order in λ, the Φ-independent values $\det\frac{\sin\Omega_0 T}{\Omega_0}$, we have

$$\left|e^{iW_{0+\phi\phi}[\Phi_1,\Phi_2]}\right| \propto \left[\det\left(\mathcal{M}^\dagger\mathcal{M}\right)\right]^{-1/4} \propto \left\{\det\left(1+\lambda^2\aleph_0^{-1/2}\aleph_1[\Delta\Phi]\aleph_0^{-1}\aleph_1[\Delta\Phi]\aleph_0^{-1/2}\right)\right\}^{-1/4}. \tag{IV.5.41}$$

The normalization is set by (IV.5.37), and in fact

$$\left|e^{iW_{0+\phi\phi}[\Phi_1,\Phi_2]}\right| = \left\{\det\left(1+\lambda^2\aleph_0^{-1/2}\aleph_1[\Delta\Phi]\aleph_0^{-1}\aleph_1[\Delta\Phi]\aleph_0^{-1/2}\right)\right\}^{-1/4}. \tag{IV.5.42}$$

For any positive matrix a^2, a straightforward analysis in the diagonal basis shows $\det(1 + a^2) \geq 1 + \mathrm{Tr}\,a^2$, so

$$\left|e^{iW_{0+\phi\phi}[\Phi_1,\Phi_2]}\right| \leq \left\{1 + \mathrm{Tr}\left(\lambda\aleph_0^{-1}\aleph_1[\Delta\Phi]\right)^2\right\}^{-1/4}. \tag{IV.5.43}$$

Using

$$\aleph_0^{-1} = \frac{1}{\Omega_0^2 \mathcal{V}(\Omega_0)} \begin{pmatrix} 1 & \cos\Omega_0 T \\ \cos\Omega_0 T & 1, \end{pmatrix} \tag{IV.5.44}$$

we have

$$\aleph_0^{-1}\aleph_1 = \frac{1}{\Omega_0^2 \mathcal{V}(\Omega_0)} \begin{pmatrix} B_1 - \cos\Omega_0 T {C_1}^{\mathrm{tr}} & -C_1 + \cos\Omega_0 T A_1 \\ \cos\Omega_0 T B_1 - {C_1}^{\mathrm{tr}} & -\cos\Omega_0 T C_1 + A_1 \end{pmatrix} \tag{IV.5.45}$$

so[3]

$$\begin{aligned}
\operatorname{Tr}\left(\lambda\aleph_0^{-1}\aleph_1\right)^2 = \operatorname{Tr}\Big[&\frac{\lambda}{\Omega_0^2\mathcal{V}(\Omega_0)}(B_1 - \cos\Omega_0 T {C_1}^{\mathrm{tr}})\frac{\lambda}{\Omega_0^2\mathcal{V}(\Omega_0)}(B_1 - \cos\Omega_0 T {C_1}^{\mathrm{tr}}) \\
&+ \frac{\lambda}{\Omega_0^2\mathcal{V}(\Omega_0)}(-C_1 + \cos\Omega_0 T A_1)\frac{\lambda}{\Omega_0^2\mathcal{V}(\Omega_0)}(\cos\Omega_0 T B_1 - {C_1}^{\mathrm{tr}}) \\
&+ \frac{\lambda}{\Omega_0^2\mathcal{V}(\Omega_0)}(\cos\Omega_0 T B_1 - {C_1}^{\mathrm{tr}})\frac{\lambda}{\Omega_0^2\mathcal{V}(\Omega_0)}(-C_1 + \cos\Omega_0 T A_1) \\
&+ \frac{\lambda}{\Omega_0^2\mathcal{V}(\Omega_0)}(-\cos\Omega_0 T C_1 + A_1)\frac{\lambda}{\Omega_0^2\mathcal{V}(\Omega_0)}(-\cos\Omega_0 T C_1 + A_1)\Big] \\
= \sum_{\mathbf{M}_1,\mathbf{M}_2\in\mathcal{S}} &\frac{\lambda^2}{k_{\mathbf{M}_1}^2\mathcal{V}(k_{\mathbf{M}_1})k_{\mathbf{M}_2}^2\mathcal{V}(k_{\mathbf{M}_2})} \begin{pmatrix} B_{1\mathbf{M}_2\mathbf{M}_1} \\ A_{1\mathbf{M}_2\mathbf{M}_1} \\ C^{\mathrm{tr}}_{1\mathbf{M}_2\mathbf{M}_1} \\ C_{1\mathbf{M}_2\mathbf{M}_1} \end{pmatrix}^{\mathrm{tr}} \\
\times &\begin{pmatrix} 1 & \cos k_{\mathbf{M}_1}T\cos k_{\mathbf{M}_2}T & -\cos k_{\mathbf{M}_2}T & -\cos k_{\mathbf{M}_1}T \\ \cos k_{\mathbf{M}_1}T\cos k_{\mathbf{M}_2}T & 1 & -\cos k_{\mathbf{M}_1}T & -\cos k_{\mathbf{M}_2}T \\ -\cos k_{\mathbf{M}_2}T & -\cos k_{\mathbf{M}_1}T & 1 & \cos k_{\mathbf{M}_1}T\cos k_{\mathbf{M}_2}T \\ -\cos k_{\mathbf{M}_1}T & -\cos k_{\mathbf{M}_2}T & \cos k_{\mathbf{M}_1}T\cos k_{\mathbf{M}_2}T & 1 \end{pmatrix} \\
&\times \begin{pmatrix} B_{1\mathbf{M}_1\mathbf{M}_2} \\ A_{1\mathbf{M}_1\mathbf{M}_2} \\ C_{1\mathbf{M}_1\mathbf{M}_2} \\ C^{\mathrm{tr}}_{1\mathbf{M}_1\mathbf{M}_2} \end{pmatrix}.
\end{aligned} \tag{IV.5.46}$$

If we define the shorthand $c_\pm = \cos\frac{k_\pm T}{2}$ and $s_\pm = \sin\frac{k_\pm T}{2}$ (where $k_\pm = k_{\mathbf{M}_1} \pm k_{\mathbf{M}_2}$) and note that

$$\cos k_{\mathbf{M}_1}T\cos k_{\mathbf{M}_2}T = \frac{1}{2}(\cos k_+T + \cos k_-T) = 1 - s_+^2 - s_-^2 = c_+^2 + c_-^2 - 1 \tag{IV.5.47a}$$

$$\sin k_{\mathbf{M}_1}T\sin k_{\mathbf{M}_2}T = \frac{1}{2}(-\cos k_+T + \cos k_-T) = s_+^2 - s_-^2 = -c_+^2 + c_-^2 \tag{IV.5.47b}$$

$$\cos k_{\mathbf{M}_1}T = \cos\frac{k_+ + k_-}{2}T = c_+c_- - s_+s_- \tag{IV.5.47c}$$

$$\cos k_{\mathbf{M}_2}T = \cos\frac{k_+ - k_-}{2}T = c_+c_- + s_+s_-, \tag{IV.5.47d}$$

[3] For the conversion of the range of the indices of these real matrices from $\mathcal{S}/2 \oplus \mathcal{S}/2$ to $\mathcal{S}$, see (IV.C.8).

we can use the result (IV.C.28) from Sec. IV.C.2 of Appendix IV.C, written as

$$\begin{pmatrix} B_{1\mathbf{M}_1\mathbf{M}_2} \\ A_{1\mathbf{M}_1\mathbf{M}_2} \\ C_{1\mathbf{M}_1\mathbf{M}_2} \\ C^{\mathrm{tr}}_{1\mathbf{M}_1\mathbf{M}_2} \end{pmatrix} = -\frac{k_{\mathbf{M}_1}k_{\mathbf{M}_2}}{2(s_+^2 - s_-^2)} \begin{pmatrix} c_- & -s_- & c_+ & -s_+ \\ c_- & s_- & c_+ & s_+ \\ c_+ & -s_+ & c_- & -s_- \\ c_+ & s_+ & c_- & s_- \end{pmatrix} \begin{pmatrix} \chi_- \\ \sigma_- \\ \chi_+ \\ \sigma_+ \end{pmatrix} \tag{IV.C.28$'$}$$

to say

$$\begin{aligned}
\mathrm{Tr}\left(\lambda\aleph_0^{-1}\aleph_1\right)^2 &= \sum_{\mathbf{M}_1,\mathbf{M}_2\in\mathcal{S}} \frac{\lambda^2}{k^2_{\mathbf{M}_1}\mathcal{V}(k_{\mathbf{M}_1})k^2_{\mathbf{M}_2}\mathcal{V}(k_{\mathbf{M}_2})} \begin{pmatrix} B_{1\mathbf{M}_1\mathbf{M}_2} \\ A_{1\mathbf{M}_1\mathbf{M}_2} \\ C_{1\mathbf{M}_1\mathbf{M}_2} \\ C^{\mathrm{tr}}_{1\mathbf{M}_1\mathbf{M}_2} \end{pmatrix}^{\mathrm{tr}} \\
&\times \begin{pmatrix} 1 & 1 - s_+^2 - s_-^2 & -c_+c_- - s_+s_- & -c_+c_- + s_+s_- \\ 1 - s_+^2 - s_-^2 & 1 & -c_+c_- + s_+s_- & -c_+c_- - s_+s_- \\ -c_+c_- - s_+s_- & -c_+c_- + s_+s_- & 1 & 1 - s_+^2 - s_-^2 \\ -c_+c_- + s_+s_- & -c_+c_- - s_+s_- & 1 - s_+^2 - s_-^2 & 1 \end{pmatrix} \begin{pmatrix} B_{1\mathbf{M}_1\mathbf{M}_2} \\ A_{1\mathbf{M}_1\mathbf{M}_2} \\ C_{1\mathbf{M}_1\mathbf{M}_2} \\ C^{\mathrm{tr}}_{1\mathbf{M}_1\mathbf{M}_2} \end{pmatrix} \\
&= \sum_{\mathbf{M}_1,\mathbf{M}_2\in\mathcal{S}} \frac{\lambda^2}{\mathcal{V}(k_{\mathbf{M}_1})\mathcal{V}(k_{\mathbf{M}_2})4(s_+^2 - s_-^2)^2} \begin{pmatrix} \chi_- \\ \sigma_- \\ \chi_+ \\ \sigma_+ \end{pmatrix}^{\mathrm{tr}} \begin{pmatrix} c_- & c_- & c_+ & c_+ \\ -s_- & s_- & -s_+ & s_+ \\ c_+ & c_+ & c_- & c_- \\ -s_+ & s_+ & -s_- & s_- \end{pmatrix} \\
&\times \begin{pmatrix} c_-(c_-^2 - c_+^2) & s_-(s_+^2 - s_-^2) & c_+(c_+^2 - c_-^2) & s_+(s_-^2 - s_+^2) \\ c_-(c_-^2 - c_+^2) & s_-(s_-^2 - s_+^2) & c_+(c_+^2 - c_-^2) & s_+(s_+^2 - s_-^2) \\ c_+(c_+^2 - c_-^2) & s_+(s_-^2 - s_+^2) & c_-(c_-^2 - c_+^2) & s_-(s_+^2 - s_-^2) \\ c_+(c_+^2 - c_-^2) & s_+(s_+^2 - s_-^2) & c_-(c_-^2 - c_+^2) & s_-(s_-^2 - s_+^2) \end{pmatrix} \begin{pmatrix} \chi_- \\ \sigma_- \\ \chi_+ \\ \sigma_+ \end{pmatrix} \\
&= \sum_{\mathbf{M}_1,\mathbf{M}_2\in\mathcal{S}} \frac{\lambda^2}{4\mathcal{V}(k_{\mathbf{M}_1})\mathcal{V}(k_{\mathbf{M}_2})} \left(\chi_-^2 + \sigma_-^2 + \chi_+^2 + \sigma_+^2\right)
\end{aligned} \tag{IV.5.48}$$

where

$$\chi_\pm = \int_{-T/2}^{T/2} dt[m_1(t) \mp n_1(t)]_{\mathbf{M}_1\mathbf{M}_2} \cos k_\pm t \mp m_1(t)_{\mathbf{M}_1\mathbf{M}_2} \frac{k_\pm}{k_{\mathbf{M}_1}k_{\mathbf{M}_2}} \sin 2k_\pm t \Big|_{-T/2}^{T/2} \tag{IV.C.30a}$$

and

$$\sigma_\pm = \int_{-T/2}^{T/2} dt[m_1(t) \mp n_1(t)]_{\mathbf{M}_1\mathbf{M}_2} \sin k_\pm t \pm m_1(t)_{\mathbf{M}_1\mathbf{M}_2} \frac{k_\pm}{k_{\mathbf{M}_1}k_{\mathbf{M}_2}} \cos 2k_\pm t \Big|_{-T/2}^{T/2} . \tag{IV.C.30b}$$

Expressing the influence phase in terms of the original field

Now it's time to take the result in terms of the real matrices $m_1[\Phi]$ and $\varpi_1[\Phi]$ on $\mathbb{R}^{\mathcal{S}/2\oplus\mathcal{S}/2}$ defined by (IV.5.17) and (IV.5.18), and reconstruct from them useful expressions in terms of $\{\Phi_{\mathbf{N}}\}$ and the complex matrices

$$(m_{1\mathrm{ex}})_{\mathbf{M}_1\mathbf{M}_2} = \Phi_{\mathbf{M}_1-\mathbf{M}_2} \tag{IV.5.49a}$$
$$(\varpi_{1\mathrm{ex}})_{\mathbf{M}_1\mathbf{M}_2} = k^2_{\mathbf{M}_1\mathbf{M}_2}(m_{1\mathrm{ex}})_{\mathbf{M}_1-\mathbf{M}_2} + (\ddot{m}_{1\mathrm{ex}})_{\mathbf{M}_1-\mathbf{M}_2} \tag{IV.5.49b}$$

on $\mathbb{C}^{\mathcal{S}}$ defined by (IV.5.6), as well as [*cf.* (IV.C.21)]

$$\begin{aligned}(n_{1\mathrm{ex}})_{\mathbf{M}_1\mathbf{M}_2} =& \frac{(\varpi_{1\mathrm{ex}})_{\mathbf{M}_1\mathbf{M}_2} - (\ddot{m}_{1\mathrm{ex}})_{\mathbf{M}_1\mathbf{M}_2} - (m_{1\mathrm{ex}})_{\mathbf{M}_1\mathbf{M}_2}(k^2_{\mathbf{M}_1} + k^2_{\mathbf{M}_2})}{k_{\mathbf{M}_1}k_{\mathbf{M}_2}} \\ =& \frac{(m_{1\mathrm{ex}})_{\mathbf{M}_1\mathbf{M}_2}(k^2_{\mathbf{M}_1\mathbf{M}_2} - k^2_{\mathbf{M}_1} - k^2_{\mathbf{M}_2})}{k_{\mathbf{M}_1}k_{\mathbf{M}_2}} = -(m_{1\mathrm{ex}})_{\mathbf{M}_1\mathbf{M}_2}\frac{\mathbf{k}_{\mathbf{M}_1}\cdot\mathbf{k}_{\mathbf{M}_2}}{k_{\mathbf{M}_1}k_{\mathbf{M}_2}} \\ =& -(m_{1\mathrm{ex}})_{\mathbf{M}_1\mathbf{M}_2}\cos\theta_{\mathbf{M}_1\mathbf{M}_2}.\end{aligned} \tag{IV.5.49c}$$

These are related to the real matrices m and ϖ on $\mathbb{R}^{\mathcal{S}}$ (or $\mathbb{R}^{\mathcal{S}/2\oplus\mathcal{S}/2}$) by (IV.5.14). Namely, for $\mathbf{M}_1, \mathbf{M}_2 \in \mathcal{S}/2$,

$$M^{\mathsf{UL}}_{\mathbf{M}_1\mathbf{M}_2} = M_{\mathbf{M}_1\mathbf{M}_2} = (M_{\mathrm{ex}})^{\mathrm{R}}_{\mathbf{M}_1\mathbf{M}_2} + (M_{\mathrm{ex}})^{\mathrm{R}}_{\mathbf{M}_1,-\mathbf{M}_2} \tag{IV.5.50a}$$
$$M^{\mathsf{LR}}_{\mathbf{M}_1\mathbf{M}_2} = M_{-\mathbf{M}_1,-\mathbf{M}_2} = (M_{\mathrm{ex}})^{\mathrm{R}}_{\mathbf{M}_1\mathbf{M}_2} - (M_{\mathrm{ex}})^{\mathrm{R}}_{\mathbf{M}_1,-\mathbf{M}_2} \tag{IV.5.50b}$$
$$M^{\mathsf{LL}}_{\mathbf{M}_1\mathbf{M}_2} = M_{-\mathbf{M}_1\mathbf{M}_2} = (M_{\mathrm{ex}})^{\mathrm{I}}_{\mathbf{M}_1\mathbf{M}_2} + (M_{\mathrm{ex}})^{\mathrm{I}}_{\mathbf{M}_1,-\mathbf{M}_2} \tag{IV.5.50c}$$
$$M^{\mathsf{UR}}_{\mathbf{M}_1\mathbf{M}_2} = M_{\mathbf{M}_1,-\mathbf{M}_2} = -(M_{\mathrm{ex}})^{\mathrm{I}}_{\mathbf{M}_1\mathbf{M}_2} + (M_{\mathrm{ex}})^{\mathrm{I}}_{\mathbf{M}_1,-\mathbf{M}_2}. \tag{IV.5.50d}$$

Since

$$\begin{aligned}&\sum_{\mathbf{M}_1,\mathbf{M}_2\in\mathcal{S}} (M_{\mathbf{M}_1\mathbf{M}_2})^2 \\ &= 2\sum_{\mathbf{M}_1,\mathbf{M}_2\in\mathcal{S}/2}\left\{\left[(M_{\mathrm{ex}})^{\mathrm{R}}_{\mathbf{M}_1\mathbf{M}_2}\right]^2 + \left[(M_{\mathrm{ex}})^{\mathrm{R}}_{\mathbf{M}_1,-\mathbf{M}_2}\right]^2 + \left[(M_{\mathrm{ex}})^{\mathrm{I}}_{\mathbf{M}_1\mathbf{M}_2}\right]^2 + \left[(M_{\mathrm{ex}})^{\mathrm{I}}_{\mathbf{M}_1,-\mathbf{M}_2}\right]^2\right\} \\ &= \sum_{\mathbf{M}_1,\mathbf{M}_2\in\mathcal{S}} \left|M_{\mathbf{M}_1\mathbf{M}_2}\right|^2,\end{aligned} \tag{IV.5.51}$$

we can use (IV.5.49a) and (IV.5.49c), along with $\left|\sum_i \alpha_i \beta_i^{\mathrm{R}}\right|^2 + \left|\sum_i \alpha_i \beta_i^{\mathrm{I}}\right|^2 = \frac{1}{2}\Big(|\sum_i \alpha_i \beta_i|^2$ $+ |\sum_i \alpha_i \beta_i^*|^2\Big)$, to write

$$\begin{aligned}
\mathrm{Tr}(\lambda \aleph_0^{-1} \aleph_1)^2 = & \sum_{\mathbf{M}_1, \mathbf{M}_2 \in \mathcal{S}} \frac{\lambda^2}{4\mathcal{V}(k_{\mathbf{M}_1})\mathcal{V}(k_{\mathbf{M}_2})} \\
\times \Bigg\{ & \left| \sin^2 \frac{\theta_{\mathbf{M}_1\mathbf{M}_2}}{2} \int_{-T/2}^{T/2} dt \Delta\Phi(t)_{\mathbf{M}_1 - \mathbf{M}_2} e^{ik_- t} - i \frac{k_-}{k_{\mathbf{M}_1} k_{\mathbf{M}_2}} \left[e^{i2k_- t} \Delta\Phi(t)_{\mathbf{M}_1 - \mathbf{M}_2} \right]_{-T/2}^{T/2} \right|^2 \\
& + \left| \sin^2 \frac{\theta_{\mathbf{M}_1\mathbf{M}_2}}{2} \int_{-T/2}^{T/2} dt \Delta\Phi(t)_{\mathbf{M}_1 - \mathbf{M}_2} e^{-ik_- t} + i \frac{k_-}{k_{\mathbf{M}_1} k_{\mathbf{M}_2}} \left[e^{-i2k_- t} \Delta\Phi(t)_{\mathbf{M}_1 - \mathbf{M}_2} \right]_{-T/2}^{T/2} \right|^2 \\
& + \left| \cos^2 \frac{\theta_{\mathbf{M}_1\mathbf{M}_2}}{2} \int_{-T/2}^{T/2} dt \Delta\Phi(t)_{\mathbf{M}_1 - \mathbf{M}_2} e^{ik_+ t} + i \frac{k_+}{k_{\mathbf{M}_1} k_{\mathbf{M}_2}} \left[e^{i2k_+ t} \Delta\Phi(t)_{\mathbf{M}_1 - \mathbf{M}_2} \right]_{-T/2}^{T/2} \right|^2 \\
& + \left| \cos^2 \frac{\theta_{\mathbf{M}_1\mathbf{M}_2}}{2} \int_{-T/2}^{T/2} dt \Delta\Phi(t)_{\mathbf{M}_1 - \mathbf{M}_2} e^{-ik_+ t} - i \frac{k_+}{k_{\mathbf{M}_1} k_{\mathbf{M}_2}} \left[e^{-i2k_+ t} \Delta\Phi(t)_{\mathbf{M}_1 - \mathbf{M}_2} \right]_{-T/2}^{T/2} \right|^2 \Bigg\}.
\end{aligned} \tag{IV.5.52}$$

This can be returned to continuum form by using $\Phi_{\mathbf{N}} = \varphi_{\mathbf{q_N}}(\delta^D k)^{1/2}\Theta(k_c - q_{\mathbf{N}})$ and $\lambda = \frac{\ell(\delta^D k)^{1/2}}{(2\pi)^{D/2}}$ to give

$$\begin{aligned}
\mathrm{Tr}(\lambda \aleph_0^{-1} \aleph_1)^2 = & \int_{k_1, k_2 > k_c} \frac{d^D k_1 d^D k_2 \Theta(k_c - q)}{4(2\pi)^D \mathcal{V}(k_1)\mathcal{V}(k_2)} \\
\times \Bigg\{ & \left| \sin^2 \frac{\theta_{\mathbf{k}_1\mathbf{k}_2}}{2} \int_{-T/2}^{T/2} dt \ell \Delta\varphi_{\mathbf{q}}(t) e^{ik_- t} - i \frac{k_-}{k_1 k_2} \left[e^{i2k_- t} \ell \Delta\varphi_{\mathbf{q}}(t) \right]_{-T/2}^{T/2} \right|^2 \\
& + \left| \sin^2 \frac{\theta_{\mathbf{k}_1\mathbf{k}_2}}{2} \int_{-T/2}^{T/2} dt \ell \Delta\varphi_{\mathbf{q}}(t) e^{-ik_- t} + i \frac{k_-}{k_1 k_2} \left[e^{-i2k_- t} \ell \Delta\varphi_{\mathbf{q}}(t) \right]_{-T/2}^{T/2} \right|^2 \\
& + \left| \cos^2 \frac{\theta_{\mathbf{k}_1\mathbf{k}_2}}{2} \int_{-T/2}^{T/2} dt \ell \Delta\varphi_{\mathbf{q}}(t) e^{ik_+ t} + i \frac{k_+}{k_1 k_2} \left[e^{i2k_+ t} \ell \Delta\varphi_{\mathbf{q}}(t) \right]_{-T/2}^{T/2} \right|^2 \\
& + \left| \cos^2 \frac{\theta_{\mathbf{k}_1\mathbf{k}_2}}{2} \int_{-T/2}^{T/2} dt \ell \Delta\varphi_{\mathbf{q}}(t) e^{-ik_+ t} - i \frac{k_+}{k_1 k_2} \left[e^{-i2k_+ t} \ell \Delta\varphi_{\mathbf{q}}(t) \right]_{-T/2}^{T/2} \right|^2 \Bigg\},
\end{aligned} \tag{IV.5.53}$$

where now $\mathbf{q} = \mathbf{k}_1 - \mathbf{k}_2$ and $k_\pm = k_1 \pm k_2$.

Specializing to D=3

If there are three spatial dimensions, the integration in (IV.5.53) is over the six components of $\mathbf{k}_1$ and $\mathbf{k}_2$. The integrand, however, is expressed in terms of the three components of $\mathbf{q}$ and the two amplitudes k_1 and k_2 (or equivalently, $k_\pm$). There is also a dependence on $\cos\theta_{\mathbf{k}_1\mathbf{k}_2} = \frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1 k_2}$, but that can be expressed in terms of the other five variables by

$$q^2 = k_1^2 + k_2^2 - 2k_1k_2\cos\theta_{\mathbf{k}_1\mathbf{k}_2}. \tag{IV.5.54}$$

To effect the change of variables from $\{\mathbf{k}_1, \mathbf{k}_2\}$ to $\{\mathbf{q}, k_+, k_-\}$, we first let $\overline{\mathbf{k}} = \frac{\mathbf{k}_1+\mathbf{k}_2}{2}$ and transform

$$d^3k_1 d^3k_2 = d^3q d^3\overline{k} = d^3q\overline{k}^2 d\overline{k} d\mu d\phi \tag{IV.5.55}$$

where $\theta = \cos^{-1}\mu$ and ϕ are the polar angles of $\overline{\mathbf{k}}$ relative to $\mathbf{q}$ (*i.e.*, $\overline{\mathbf{k}}\cdot\mathbf{q} = \overline{k}q\mu$). The integrand on (IV.5.53) is independent of the azimuthal angle ϕ. We want to transform the coördinates $\overline{k}$ and μ to k_+ and k_- for a given $\mathbf{q}$. We can find $\overline{k}$ by

$$\overline{k}^2 = \left(\frac{\mathbf{k}_1+\mathbf{k}_2}{2}\right)^2 = \frac{k_1^2+k_2^2+2k_1k_2\cos\theta_{\mathbf{k}_1\mathbf{k}_2}}{4} = \frac{k_1^2+k_2^2+(k_1^2+k_2^2-q^2)}{4} = \frac{k_1^2+k_2^2}{2} - \frac{q^2}{4}; \tag{IV.5.56}$$

to obtain a usable expression for μ, we consider

$$k_1^2 = \left(\overline{\mathbf{k}} + \frac{\mathbf{q}}{2}\right)^2 = \overline{k}^2 + \overline{k}q\mu + \frac{q^2}{4} \tag{IV.5.57a}$$

$$k_2^2 = \left(\overline{\mathbf{k}} - \frac{\mathbf{q}}{2}\right)^2 = \overline{k}^2 - \overline{k}q\mu + \frac{q^2}{4}, \tag{IV.5.57b}$$

so

$$\mu = \frac{k_1^2 - k_2^2}{2\overline{k}q}. \tag{IV.5.58}$$

We can convert these into expressions involving $k_\pm$ by using

$$k_1^2 + k_2^2 = \frac{k_+^2 + k_-^2}{2}, \qquad k_1k_2 = \frac{k_+^2 - k_-^2}{4}, \qquad k_1^2 - k_2^2 = k_+k_-, \tag{IV.5.59}$$

to give

$$\mu = \frac{k_+k_-}{2\overline{k}q} \tag{IV.5.60a}$$

$$\overline{k}^2 = \frac{k_+^2 + k_-^2 - q^2}{4}. \tag{IV.5.60b}$$

The Jacobian matrix corresponding to this transformation is

$$\begin{pmatrix} d\overline{k} \\ d\mu \end{pmatrix} = J\begin{pmatrix} dk_+ \\ dk_- \end{pmatrix} = \begin{pmatrix} \frac{k_+}{4\overline{k}} & \frac{k_-}{4\overline{k}} \\ \frac{k_-}{2\overline{k}q} - \frac{k_+k_-}{2\overline{k}^2q}\left(\frac{k_+}{4\overline{k}}\right) & \frac{k_+}{2\overline{k}q} - \frac{k_+k_-}{2\overline{k}^2q}\left(\frac{k_-}{4\overline{k}}\right) \end{pmatrix}\begin{pmatrix} dk_+ \\ dk_- \end{pmatrix}, \tag{IV.5.61}$$

so the Jacobian determinant is

$$|\det J| = \frac{k_+^2 - k_-^2}{8\overline{k}^2 q} \tag{IV.5.62}$$

(since $k_+^2 - k_-^2 = 4k_1k_2 > 0$). Thus the measure becomes

$$d^3k_1 d^3k_2 = d^3q\overline{k}^2 d\overline{k} d\mu d\phi = \frac{(k_+^2 - k_-^2)d^3q dk_+ dk_- d\phi}{8q}; \tag{IV.5.63}$$

since the integrand is independent of ϕ, we will assume that integral has already been performed and replace it with 2π.

Now that we've converted the measure, we want to consider the limits of integration on the five variables $\{\mathbf{q}, k_+, k_-\}$. There are two sources for this limit. First, there are the inherent geometrical restrictions involved in what values $k_\pm$ take as $\mathbf{k}_1$ and $\mathbf{k}_2$ each range over $\mathbb{R}^3$. Second, the conditions $k_1, k_2 \geq k_c \geq q$ should then be added. We consider the first set of constraints on the region of integration first. The condition of interest is

$$1 \geq \cos\theta_{\mathbf{k}_1\mathbf{k}_2} \geq -1; \tag{IV.5.64}$$

since

$$\cos\theta_{\mathbf{k}_1\mathbf{k}_2} = \frac{k_1^2 + k_2^2 - q^2}{2k_1k_2} = \frac{k_+^2 + k_-^2 - 2q^2}{k_+^2 - k_-^2}, \tag{IV.5.65}$$

this means

$$k_+^2 + k_-^2 - 2q^2 \leq k_+^2 - k_-^2 \geq 2q^2 - k_+^2 - k_-^2, \tag{IV.5.66}$$

from which we get the restrictions

$$k_-^2 \leq q^2 \leq k_+^2. \tag{IV.5.67}$$

By its construction, k_+ is inherently positive, but k_- is not. This makes the geometric limits of integration

$$\int d^3k_1 d^3k_2 = 2\pi \int_0^\infty dq \iint q^2 d^2\Omega_{\hat{q}} \int_q^\infty dk_+ \int_{-q}^{q} dk_- \frac{k_+^2 - k_-^2}{8q}. \tag{IV.5.68}$$

Moving to the second set of restrictions, $q \leq k_c$ is easy to add, while the conditions

$$k_+ + k_- = 2k_1 \geq 2k_c \tag{IV.5.69a}$$
$$k_+ - k_- = 2k_2 \geq 2k_c \tag{IV.5.69b}$$

translate to either

$$k_+ \geq 2k_c, \qquad |k_-| \leq k_+ - 2k_c \tag{IV.5.70a}$$

or

$$\text{any } k_-, \qquad k_+ \geq 2k_c + |k_-|\,. \tag{IV.5.70b}$$

Combining the two sets of limits gives (see Fig. IV.2)

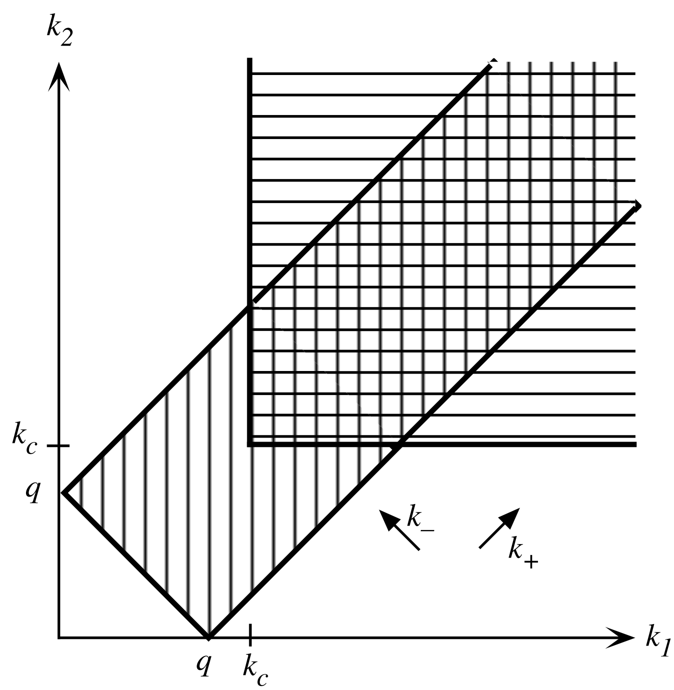


Figure IV.2: The regions of integration for (IV.5.53). The inherent geometrical restrictions $k_+ = k_1 + k_2 \geq q$ and $|k_-| = |k_1 - k_2| \leq q$ limit us to the region shaded vertically, while the additional requirement that $k_1, k_2 \geq k_c$ requires that the mode be in the region shaded horizontally. Their intersection gives the region of integration for (IV.5.53).

$$\begin{aligned}\int\limits_{k_1,k_2>k_c>|\mathbf{k}_1-\mathbf{k}_2|} d^3k_1 d^3k_2 =& 2\pi \int\limits_{\text{limits}} d^3q\,dk_+dk_- \frac{k_+^2-k_-^2}{8q} \\ =& 2\pi\int_0^\infty dq \iint q^2 d^2\Omega_{\hat{q}} \int_{-q}^{q} dk_- \int_{2k_c+|k_-|}^{\infty} dk_+ \frac{k_+^2-k_-^2}{8q}\end{aligned} \tag{IV.5.71a}$$

$$= 2\pi\int_0^\infty dq \iint q^2 d^2\Omega_{\hat{q}} \left(\int_{2k_c}^{2k_c+q} dk_+ \int_{-(k_+-2k_c)}^{k_+-2k_c} dk_- + \int_{2k_c+q}^{\infty} dk_+ \int_{-q}^{q} dk_-\right) \frac{k_+^2-k_-^2}{8q}. \tag{IV.5.71b}$$

(The first of these two is more convenient to work with.) This means that in the three-dimensional case, (IV.5.53) becomes

$$\begin{aligned}\mathrm{Tr}(\lambda\aleph_0^{-1}\aleph_1)^2 = \int\limits_{\text{limits}} & \frac{d^3q\,dk_+dk_-(k_+^2-k_-^2)}{(2\pi)^2 32q\mathcal{V}\left(\frac{k_++k_-}{2}\right)\mathcal{V}\left(\frac{k_+-k_-}{2}\right)} \\ \times\Bigg\{ & \left|\frac{q^2-k_-^2}{k_+^2-k_-^2}\int_{-T/2}^{T/2} dt\ell\Delta\varphi_{\mathbf{q}}(t)e^{ik_-t} - i\frac{4k_-}{k_+^2-k_-^2}\left[e^{i2k_-t}\ell\Delta\varphi_{\mathbf{q}}(t)\right]_{-T/2}^{T/2}\right|^2 \\ & + \left|\frac{q^2-k_-^2}{k_+^2-k_-^2}\int_{-T/2}^{T/2} dt\ell\Delta\varphi_{\mathbf{q}}(t)e^{-ik_-t} + i\frac{4k_-}{k_+^2-k_-^2}\left[e^{-i2k_-t}\ell\Delta\varphi_{\mathbf{q}}(t)\right]_{-T/2}^{T/2}\right|^2 \\ & + \left|\frac{k_+^2-q^2}{k_+^2-k_-^2}\int_{-T/2}^{T/2} dt\ell\Delta\varphi_{\mathbf{q}}(t)e^{ik_+t} + i\frac{4k_+}{k_+^2-k_-^2}\left[e^{i2k_+t}\ell\Delta\varphi_{\mathbf{q}}(t)\right]_{-T/2}^{T/2}\right|^2 \\ & + \left|\frac{k_+^2-q^2}{k_+^2-k_-^2}\int_{-T/2}^{T/2} dt\ell\Delta\varphi_{\mathbf{q}}(t)e^{-ik_+t} - i\frac{4k_+}{k_+^2-k_-^2}\left[e^{-i2k_+t}\ell\Delta\varphi_{\mathbf{q}}(t)\right]_{-T/2}^{T/2}\right|^2\Bigg\};\end{aligned} \tag{IV.5.72}$$

noting that

$$\frac{1}{\mathcal{V}(k_1)\mathcal{V}(k_2)} = \frac{k_1k_2}{4}\coth\frac{\beta k_1}{2}\coth\frac{\beta k_2}{2} = \frac{k_+^2-k_-^2}{16}\coth\frac{\beta k_1}{2}\coth\frac{\beta k_2}{2}, \tag{IV.5.73}$$

we have

$$\mathrm{Tr}(\lambda\aleph_0^{-1}\aleph_1)^2 = \int_0^\infty dq \iint q^2 d^2\Omega_{\hat{q}} \int_{-q}^{q} dk_- \int_{2k_c+|k_-|}^{\infty} dk_+ \frac{\coth\beta\frac{k_++k_-}{4}\coth\beta\frac{k_+-k_-}{4}}{(2\pi)^2 512q}$$
$$\times\Bigg\{ \left|(q^2-k_-^2)\int_{-T/2}^{T/2} dt\ell\Delta\varphi_{\mathbf{q}}(t)e^{ik_-t} - i4k_-\left[e^{i2k_-t}\ell\Delta\varphi_{\mathbf{q}}(t)\right]_{-T/2}^{T/2}\right|^2$$
$$+\left|(q^2-k_-^2)\int_{-T/2}^{T/2} dt\ell\Delta\varphi_{\mathbf{q}}(t)e^{-ik_-t} + i4k_-\left[e^{-i2k_-t}\ell\Delta\varphi_{\mathbf{q}}(t)\right]_{-T/2}^{T/2}\right|^2$$
$$+\left|(k_+^2-q^2)\int_{-T/2}^{T/2} dt\ell\Delta\varphi_{\mathbf{q}}(t)e^{ik_+t} + i4k_+\left[e^{i2k_+t}\ell\Delta\varphi_{\mathbf{q}}(t)\right]_{-T/2}^{T/2}\right|^2$$
$$+\left|(k_+^2-q^2)\int_{-T/2}^{T/2} dt\ell\Delta\varphi_{\mathbf{q}}(t)e^{-ik_+t} - i4k_+\left[e^{-i2k_+t}\ell\Delta\varphi_{\mathbf{q}}(t)\right]_{-T/2}^{T/2}\right|^2\Bigg\}. \tag{IV.5.74}$$

IV.6 The full action

Now we need to consider the effects of adding the terms S_ϕ and $S_{\phi\phi\phi}$ back into the action, and determine what effect, if any, this has on the influence phase (IV.5.43).

IV.6.1 The linear terms

The effect of the linear term S_ϕ can, as usual, be elucidated by completing the square, as shown in this section.

We define the "all-but-cubic" lagrangian

$$L_3[\phi,\Phi] = L_{0+\phi\phi}[\phi,\Phi] + \lambda L_\phi[\phi,\Phi] \tag{IV.6.1}$$

by adding to the quadratic action considered in Sec. IV.5 the linear terms [*cf.* (IV.4.8c)]

$$\lambda L_\phi[\phi,\Phi] = \lambda\sum_{\mathbf{M}\in\mathcal{S}}\left(-\phi^*_{\mathbf{M}}\widetilde{x}_{\mathbf{M}} + \phi_{\mathbf{M}}\widetilde{y}_{\mathbf{M}}\right) \tag{IV.6.2}$$

where[4] [*cf.* (IV.4.8c)

$$\widetilde{x}_{\mathbf{M}} = -\frac{\lambda}{2}\left(q^2_{\mathbf{M-N,N}}\Phi_{\mathbf{M-N}}\Phi_{\mathbf{N}} - \dot{\Phi}_{\mathbf{M-N}}\dot{\Phi}_{\mathbf{N}}\right) \tag{IV.6.3a}$$
$$\widetilde{y}_{\mathbf{M}} = -\frac{\lambda}{2}\left(\dot{\Phi}_{\mathbf{M-N}}\Phi_{\mathbf{N}} - \Phi_{\mathbf{M-N}}\dot{\Phi}_{\mathbf{N}}\right), \tag{IV.6.3b}$$

with $q^2_{\mathbf{M-N,N}}$ defined analogously [*cf.* (IV.5.3)] to $k^2_{\mathbf{M}_1\mathbf{M}_2}$:

$$q^2_{\mathbf{M-N,N}} = -\mathbf{k}_{-\mathbf{M}}\cdot\mathbf{q}_{\mathbf{M-N}} - \mathbf{q}_{\mathbf{M-N}}\cdot\mathbf{q}_{\mathbf{N}} - \mathbf{q}_{\mathbf{N}}\cdot\mathbf{k}_{-\mathbf{M}} = \mathbf{k}^2_{\mathbf{M}} + \mathbf{q}^2_{\mathbf{N}} - \mathbf{k}_{\mathbf{M}}\cdot\mathbf{q}_{\mathbf{N}}. \tag{IV.6.4}$$

[4] recall that $\Phi_{\mathbf{M-N}} = 0$ when $\mathbf{M-N}\notin\mathcal{L}$

The reality condition $\Phi_{-\mathbf{N}} = \Phi^*_{\mathbf{N}}$ forces $\widetilde{x}_{-\mathbf{M}} = \widetilde{x}^*_{\mathbf{M}}$ and $\widetilde{y}_{-\mathbf{M}} = \widetilde{y}^*_{\mathbf{M}}$, so we can use the identity

$$v^{\mathrm{ex}\dagger} w^{\mathrm{ex}} = \sum_{\mathbf{M}\in\mathcal{S}} v^*_{\mathbf{M}} w_{\mathbf{M}} = \sum_{\mathbf{M}\in\mathcal{S}} \left(v^{\mathrm{R}}_{\mathbf{M}} w^{\mathrm{R}}_{\mathbf{M}} + v^{\mathrm{I}}_{\mathbf{M}} w^{\mathrm{I}}_{\mathbf{M}}\right) = 2 \sum_{\mathbf{M}\in\mathcal{S}/2} \left(v^{\mathrm{R}}_{\mathbf{M}} w^{\mathrm{R}}_{\mathbf{M}} + v^{\mathrm{I}}_{\mathbf{M}} w^{\mathrm{I}}_{\mathbf{M}}\right) = v^{\mathrm{tr}} w, \tag{IV.6.5}$$

where v and w are vectors in $\mathbb{R}^{\mathcal{S}/2\oplus\mathcal{S}/2}$ defined as in (IV.5.16), to write

$$\lambda L_\phi = -\phi^{\mathrm{tr}}\widetilde{x} + \dot{\phi}^{\mathrm{tr}}\widetilde{y}; \tag{IV.6.6}$$

by integrating by parts, we can also write this, for arbitrary $z(t)$ as

$$\lambda L_\phi = -\phi^{\mathrm{tr}} x + \dot{\phi}^{\mathrm{tr}} y + \frac{d}{dt}\left(\phi^{\mathrm{tr}} z\right), \tag{IV.6.7}$$

where

$$x = \widetilde{x} + \dot{z} \tag{IV.6.8a}$$

$$y = \widetilde{y} - z \tag{IV.6.8b}$$

to give

$$L_3[\phi,\Phi] = \frac{1}{2}\left[\dot{\phi}^{\mathrm{tr}} m\dot{\phi} + 2\dot{\phi}^{\mathrm{tr}} y - \phi^{\mathrm{tr}}\varpi\phi - 2\phi^{\mathrm{tr}} x + \frac{d}{dt}\left(\phi^{\mathrm{tr}}\dot{m}\phi + 2\phi^{\mathrm{tr}} z\right)\right]. \tag{IV.6.9}$$

Comparing this to

$$\begin{aligned} L_{0+\phi\phi}[\phi + \varpi^{-1}x] = \frac{1}{2}\Bigg[&\dot{\phi}^{\mathrm{tr}} m\dot{\phi} + 2\dot{\phi}^{\mathrm{tr}} m\frac{d}{dt}(\varpi^{-1}x) + \frac{d}{dt}(x^{\mathrm{tr}}\varpi^{-1}) m\frac{d}{dt}(\varpi^{-1}x) \\ &-\phi^{\mathrm{tr}}\varpi\phi - 2\phi^{\mathrm{tr}} x + x^{\mathrm{tr}}\varpi^{-1}x + \frac{d}{dt}(\phi^{\mathrm{tr}}\dot{m}\phi + \phi^{\mathrm{tr}}\dot{m}\varpi^{-1}x + x^{\mathrm{tr}}\varpi^{-1}\dot{m}\varpi^{-1}x)\Bigg], \end{aligned} \tag{IV.6.10}$$

we see that if

$$y = m\frac{d}{dt}(\varpi^{-1}x) \tag{IV.6.11}$$

this becomes

$$\begin{aligned} &L_{0+\phi\phi}[\phi + \varpi^{-1}x] \\ &= L_3[\phi,\Phi] + \frac{y^{\mathrm{tr}} m^{-1} y}{2} - \frac{x^{\mathrm{tr}}\varpi^{-1}x}{2} + \frac{d}{dt}\left[\phi^{\mathrm{tr}}(\dot{m}\varpi^{-1}x - z)\right] + \frac{d}{dt}\frac{x^{\mathrm{tr}}\varpi^{-1}\dot{m}\varpi^{-1}x}{2} \end{aligned} \tag{IV.6.12}$$

The condition (IV.6.11) is equivalent to the second order inhomogeneous ODE

$$\frac{d}{dt}(\varpi^{-1}\dot{z}) + m^{-1}z = m^{-1}\widetilde{y} - \frac{d}{dt}(\varpi^{-1}\widetilde{x}). \tag{IV.6.13}$$

A particular $\Phi(t)$ generates $\widetilde{x}(t)$ and $\widetilde{y}(t)$ via (IV.6.3), and for that source term, we can solve (IV.6.13), with the freedom to fix two boundary conditions which are functions of z', z'', $(\dot{z})'$ and $(\dot{z})''$.

We can use this expression for L_3 to express $\mathcal{K}_3$, the propagator for S_3, in terms of $\mathcal{K}_{0+\phi\phi}$ as

$$\begin{aligned}\mathcal{K}_3(\phi''|\phi';\Phi) =& \int\limits_{\phi''\phi'} \mathcal{D}\phi e^{iS_3[\phi,\Phi]}\\ &=e^{\frac{i}{2}\int_{-T/2}^{T/2} dt(x^{\mathrm{tr}}\varpi^{-1}x - y^{\mathrm{tr}}m^{-1}y)} e^{i\phi^{\mathrm{tr}}(z-\dot{m}\varpi^{-1}x)|_{-T/2}^{T/2}} \int\limits_{\phi''\phi'} \mathcal{D}\phi e^{iS_{0+\phi\phi}[\phi+\varpi^{-1}x,\Phi]}.\end{aligned} \tag{IV.6.14}$$

Recognizing the last expression as $\mathcal{K}_{0+\phi\phi}(\phi'' + \{\varpi^{-1}x\}''|\phi' + \{\varpi^{-1}x\}';\Phi]$, we have

$$\mathcal{K}_3(\phi''|\phi';\Phi] = \mathcal{K}_{0+\phi\phi}(\phi''|\phi';\Phi] \exp\left\{ i\begin{pmatrix}\phi''\\ \phi'\end{pmatrix}^{\mathrm{tr}} \mathcal{X}[\Phi]\right\} e^{i\psi[\Phi]}, \tag{IV.6.15}$$

where

$$\begin{aligned}\psi[\Phi] =&\frac{1}{2}\int_{-T/2}^{T/2} dt(x^{\mathrm{tr}}\varpi^{-1}x - y^{\mathrm{tr}}m^{-1}y)\\ &+\frac{1}{2}\begin{pmatrix}(\varpi^{-1}x)''\\ (\varpi^{-1}x)'\end{pmatrix}^{\mathrm{tr}} \begin{pmatrix}B[\Phi]-\dot{m}'' & -C[\Phi]\\ -C[\Phi]^{\mathrm{tr}} & A[\Phi]+\dot{m}'\end{pmatrix} \begin{pmatrix}(\varpi^{-1}x)''\\ (\varpi^{-1}x)'\end{pmatrix}\end{aligned} \tag{IV.6.16}$$

is a real phase, and

$$\mathcal{X}(\Phi) = \begin{pmatrix} z'' - (\dot{m}\varpi^{-1}x)''\\ -z'' + (\dot{m}\varpi^{-1}x)'\end{pmatrix} + \begin{pmatrix}B[\Phi] & -C[\Phi]\\ -C[\Phi]^{\mathrm{tr}} & A[\Phi]\end{pmatrix}\begin{pmatrix}(\varpi^{-1}x)''\\ (\varpi^{-1}x)'\end{pmatrix}. \tag{IV.6.17}$$

By substituting for x using (IV.6.8a), we see that $\mathcal{X}[\Phi] = 0$ is just a pair of first order boundary conditions on $z(t)$, and so we can choose the solution to (IV.6.13) to obey them, leaving

$$\mathcal{K}_3(\phi''|\phi';\Phi] = \mathcal{K}_{0+\phi\phi}(\phi''|\phi';\Phi]e^{i\psi[\Phi]}. \tag{IV.6.18}$$

Proceeding along the same lines as (IV.5.20), we find that

$$e^{iW_3[\Phi_1,\Phi_2]} = e^{iW_{0+\phi\phi}[\Phi_1,\Phi_2]}e^{i(\psi[\Phi_1]-\psi[\Phi_2])}. \tag{IV.6.19}$$

Since $\psi[\Phi]$ is real, and it is the imaginary part of W which imposes decoherence,

$$\left|e^{iW_3[\Phi_1,\Phi_2]}\right| = \left|e^{iW_{0+\phi\phi}[\Phi_1,\Phi_2]}\right| \tag{IV.6.20}$$

and adding in the linear terms does not change the result (IV.5.43).

IV.6.2 The cubic terms

Now we are ready to consider the full action

$$S_E[\phi,\Phi] = S_3[\phi,\Phi] + S_{\phi\phi\phi}[\phi] \tag{IV.6.21}$$

including the cubic terms from

$$L_{\phi\phi\phi}(\phi) = -\frac{1}{2}\sum_{\mathbf{M}_1,\mathbf{M}_2,\mathbf{M}_3\in\mathcal{S}} \delta_{\mathbf{M}_1+\mathbf{M}_2+\mathbf{M}_3,\mathbf{0}} \left(\phi_{\mathbf{M}_1}\dot{\phi}_{\mathbf{M}_2}\dot{\phi}_{\mathbf{M}_3} - \frac{\mathbf{k}^2_{\mathbf{M}_1}+\mathbf{k}^2_{\mathbf{M}_2}+\mathbf{k}^2_{\mathbf{M}_3}}{6}\phi_{\mathbf{M}_1}\phi_{\mathbf{M}_2}\phi_{\mathbf{M}_3}\right). \tag{IV.6.22}$$

Here we need to resort to using a generating functional

$$Z[J_1,J_2,\Phi_1,\Phi_2] = \int \mathcal{D}\phi_1\mathcal{D}\phi_2\rho_\phi(\phi_1',\phi_2')\delta(\phi_2''-\phi_1'') \times \exp\left\{i\left(S_3[\phi_1,\Phi_1]-S_3[\phi_2,\Phi_2]+\int_{-T/2}^{T/2} dt[{\phi_1}^{\mathrm{tr}}J_1-{\phi_2}^{\mathrm{tr}}J_2]\right)\right\} \tag{IV.6.23a}$$

and expressing[5]

$$e^{iW[\Phi_1,\Phi_2]} = \exp\left(i\lambda S_{\phi\phi\phi}\left[\frac{1}{i}\frac{\mathcal{D}}{\mathcal{D}J_1},\Phi_1\right] - i\lambda S_{\phi\phi\phi}\left[-\frac{1}{i}\frac{\mathcal{D}}{\mathcal{D}J_2},\Phi_2\right]\right) Z[J_1,J_2,\Phi_1,\Phi_2]. \tag{IV.6.23b}$$

The wrong way to complete the square

The "propagators" involved in constructing the generating functional

$$Z[J_1,J_2,\Phi_1,\Phi_2] = \int d\phi_1' d\phi_2' d\phi'' \rho_\phi(\phi_1',\phi_2')\mathcal{K}_Z(\phi''|\phi_1';\Phi_1,J_1]\mathcal{K}_Z^*(\phi''|\phi_2';\Phi_2,J_2] \tag{IV.6.24}$$

are of the form

$$\mathcal{K}_Z(\phi''|\phi';\Phi,J] = \int\limits_{\phi''\phi'} \mathcal{D}\phi e^{i\left(S_3[\phi,\Phi]+\int_{-T/2}^{T/2} dt\phi^{\mathrm{tr}}J\right)} \tag{IV.6.25}$$

and involve the modified "lagrangian"

$$L_Z[\phi,\Phi,J] = L_3[\phi,\Phi] + J^{\mathrm{tr}}\phi = \frac{1}{2}\left[\dot{\phi}^{\mathrm{tr}}m\dot{\phi} + 2\dot{\phi}^{\mathrm{tr}}\widetilde{y} - \phi^{\mathrm{tr}}\varpi\phi - 2\phi^{\mathrm{tr}}(\widetilde{x}-J) + \frac{d}{dt}\left(\phi^{\mathrm{tr}}\dot{m}\phi\right)\right]; \tag{IV.6.26}$$

We might try to carry out the same completion of the square as was done in Sec. IV.6.1, getting the form (IV.6.7), where now

$$x = \widetilde{x} + \dot{z} - J. \tag{IV.6.8a$'$}$$

The differential equation for z becomes

$$\frac{d}{dt}(\varpi^{-1}\dot{z}) + m^{-1}z = m^{-1}\widetilde{y} - \frac{d}{dt}[\varpi^{-1}(\widetilde{x}-J)]; \tag{IV.6.13$'$}$$

[5] The choice of sign of J_2 may seem unusual, but it allows us to write the argument of the exponential in (IV.6.23a) as $S_Z[\phi_1,\Phi_1,J_1]-S_Z[\phi_2,\Phi_2,J_2]$ rather than $S_Z[\phi_1,\Phi_1,J_1]-S_Z[\phi_2,\Phi_2,-J_2]$.

once again, we would find an expression like

$$\mathcal{K}_Z(\phi''|\phi';\Phi,J)=\mathcal{K}_{0+\phi\phi}(\phi''|\phi';\Phi)\exp\left\{i\begin{pmatrix}\phi''\\ \phi'\end{pmatrix}^{\mathrm{tr}}\mathcal{X}[\Phi,J]\right\}e^{i\psi[\Phi,J]}, \tag{IV.6.15$'$}$$

with the expressions (IV.6.16) and (IV.6.17) for $\psi[\Phi,J]$ and $\mathcal{X}[\Phi,J]$ in terms of x still holding. Again, the boundary conditions on (IV.6.15$'$) would allow us to set $\mathcal{X}[\Phi,J]=0$, leaving

$$\mathcal{K}_Z(\phi''|\phi';\Phi,J)=\mathcal{K}_{0+\phi\phi}(\phi''|\phi';\Phi)e^{i\psi[\Phi,J]}. \tag{IV.6.18$'$}$$

However, this form is not convenient, even if we insert $x=\widetilde{x}+\dot{z}-J$, since the expression would depend on J not only explicitly, but also implicitly via the solution $z[\Phi,J]$ to (IV.6.13$'$).

The correct way to complete the square

Since we cannot fruitfully complete the square for the J-terms in the way we did for L_ϕ in Sec. IV.6.1, let us instead combine L_ϕ with the J-terms by integrating by parts until $y=0$, *i.e.*,

$$x=\widetilde{x}+\dot{\widetilde{y}} \tag{IV.6.27a}$$
$$z=\widetilde{y} \tag{IV.6.27b}$$

so that

$$L_Z[\phi,\Phi,J]=\frac{1}{2}\left[\dot{\phi}^{\mathrm{tr}}m\dot{\phi}-\phi^{\mathrm{tr}}\varpi\phi+2\phi^{\mathrm{tr}}\widetilde{J}+\frac{d}{dt}\left(\phi^{\mathrm{tr}}\dot{m}\phi+2\phi^{\mathrm{tr}}\widetilde{y}\right)\right] \tag{IV.6.28}$$

where

$$\widetilde{J}=J-\widetilde{x}+\dot{\widetilde{y}}. \tag{IV.6.29}$$

First, we define a Green's function $G(t,t_1)$ (implicitly dependent upon Φ) satisfying

$$\left[\partial_t m(t)\partial_t+\varpi(t)\right]G(t,t_1)=\delta(t-t_1), \tag{IV.6.30}$$

so that

$$(G\circ J)(t)=\int_{-T/2}^{T/2}dt_1G(t,t_1)J(t_1) \tag{IV.6.31}$$

obeys

$$\left(\partial_t m\partial_t+\varpi\right)G\circ J=J. \tag{IV.6.32}$$

We can construct this perturbatively, with the lowest order term being

$$G_0(t,t_1)=\frac{\sin\Omega_0\left|t-t_1\right|}{2\Omega_0} \tag{IV.6.33}$$

Then we can complete the square by calculating

$$\begin{aligned}
&L_{0+\phi\phi}[\phi - G\circ\widetilde{J},\Phi] \\
&\quad = L_{0+\phi\phi}[\phi,\Phi] + L_{0+\phi\phi}[G\circ\widetilde{J},\Phi] - \left(\dot{\phi}^{\mathrm{tr}} m\dot{G}\circ\widetilde{J} - \phi^{\mathrm{tr}}\varpi G\circ\widetilde{J}\right) - \frac{d}{dt}\left(\phi^{\mathrm{tr}}\dot{m}G\circ\widetilde{J}\right) \\
&\quad = L_{0+\phi\phi}[\phi,\Phi] + L_{0+\phi\phi}[G\circ\widetilde{J},\Phi] + \phi^{\mathrm{tr}}\left(\partial_t m\partial_t + \varpi\right)G\circ\widetilde{J} - \frac{d}{dt}\left[\phi^{\mathrm{tr}}\partial_t(mG\circ\widetilde{J})\right] \\
&\quad = L_Z[\phi,\Phi,J] + L_{0+\phi\phi}[G\circ\widetilde{J},\Phi] - \frac{d}{dt}\left\{\phi^{\mathrm{tr}}[\partial_t(mG\circ\widetilde{J}) + \widetilde{y}]\right\}. \qquad \text{(IV.6.34)}
\end{aligned}$$

This gives

$$\begin{aligned}
&\mathcal{K}_Z(\phi''|\phi';\Phi,J) \\
&= \int_{\phi''\phi'} \mathcal{D}\phi \exp\left(i\left\{S_{0+\phi\phi}[\phi - G\circ\widetilde{J},\Phi] - S_{0+\phi\phi}[G\circ\widetilde{J},\Phi] + \phi^{\mathrm{tr}}[\partial_t(mG\circ\widetilde{J}) + \widetilde{y}]\Big|_{-T/2}^{T/2}\right\}\right) \\
&\qquad = \mathcal{K}_{0+\phi\phi}\left[\phi'' - (G\circ\widetilde{J})''|\phi' - (G\circ\widetilde{J})'\right] \\
&\qquad\quad \times \exp\left(i\left\{-S_{0+\phi\phi}[G\circ\widetilde{J},\Phi] + \phi^{\mathrm{tr}}[\partial_t(mG\circ\widetilde{J}) + \widetilde{y}]\Big|_{-T/2}^{T/2}\right\}\right). \qquad \text{(IV.6.35)}
\end{aligned}$$

The generating functional is thus

$$\begin{aligned}
&Z[J_1,J_2,\Phi_1,\Phi_2] = \int \frac{d\phi_1' d\phi_2' d\phi'' \rho_\phi(\phi_1',\phi_2')}{\sqrt{\det(2\pi\mathfrak{C}[\Phi_1])\det(2\pi\mathfrak{C}[\Phi_2])}} \\
&\times\exp\left\{\frac{i}{2}\begin{pmatrix}\phi'' - (G_1\circ\widetilde{J}_1)'' \\ \phi'' - (G_2\circ\widetilde{J}_2)'' \\ \phi_1' - (G_1\circ\widetilde{J}_1)' \\ \phi_2' - (G_2\circ\widetilde{J}_2)'\end{pmatrix}^{\mathrm{tr}}
\begin{pmatrix} B[\Phi_1] & 0 & -C[\Phi_1] & 0 \\ 0 & -B[\Phi_2] & 0 & C[\Phi_2] \\ -C[\Phi_1]^{\mathrm{tr}} & 0 & A[\Phi_1] & 0 \\ 0 & C[\Phi_2]^{\mathrm{tr}} & 0 & -A[\Phi_2]\end{pmatrix}
\begin{pmatrix}\phi'' - (G_1\circ\widetilde{J}_1)'' \\ \phi'' - (G_2\circ\widetilde{J}_2)'' \\ \phi_1' - (G_1\circ\widetilde{J}_1)' \\ \phi_2' - (G_2\circ\widetilde{J}_2)'\end{pmatrix}\right. \\
&\left. + i\begin{pmatrix}\phi'' \\ \phi'' \\ \phi_1' \\ \phi_2'\end{pmatrix}^{\mathrm{tr}}
\begin{pmatrix}\left[\partial_t(m_1G_1\circ\widetilde{J}_1) + \widetilde{y}_1\right]'' \\ -\left[\partial_t(m_2G_2\circ\widetilde{J}_2) + \widetilde{y}_2\right]'' \\ -\left[\partial_t(m_1G_1\circ\widetilde{J}_1) + \widetilde{y}_1\right]' \\ \left[\partial_t(m_2G_2\circ\widetilde{J}_2) + \widetilde{y}_2\right]'\end{pmatrix}
- iS_{0+\phi\phi}[G_1\circ\widetilde{J}_1,\Phi_1] + iS_{0+\phi\phi}[G_2\circ\widetilde{J}_2,\Phi_2]\right\};
\end{aligned}$$

(IV.6.36)

inserting the form of ρ_ϕ from (IV.5.27), and making the transformation (*cf.* (IV.5.28))

$$\begin{pmatrix}\phi_1'' \\ \phi_2'' \\ \phi_1' \\ \phi_2'\end{pmatrix} = \begin{pmatrix}1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1\end{pmatrix}\begin{pmatrix}\Delta\phi''/2 \\ \overline{\phi''} \\ \overline{\phi'} \\ \Delta\phi'/2\end{pmatrix}, \qquad \text{(IV.6.37)}$$

we have

$$
\begin{aligned}
Z[J_1,J_2,\Phi_1,\Phi_2] \propto \int & \frac{d\overline{\phi'}d\Delta\phi' d\phi''}{\sqrt{\det(2\pi\mathfrak{C}[\Phi_1])\det(2\pi\mathfrak{C}[\Phi_2])}} \\
\times \exp\Bigg\{ & \frac{i}{2}\begin{pmatrix} -\Delta(G\circ\widetilde{J})''/2 \\ \phi'' - \overline{(G\circ\widetilde{J})''} \\ \overline{\phi'} - \overline{(G\circ\widetilde{J})'} \\ \Delta\phi'/2 - \Delta(G\circ\widetilde{J})'/2 \end{pmatrix}^{\mathrm{tr}} \begin{pmatrix} \mathcal{B}_- & \mathcal{B}_+ & -\mathcal{C}_+ & -\mathcal{C}_- \\ \mathcal{B}_+ & \mathcal{B}_- & -\mathcal{C}_- & -\mathcal{C}_+ \\ -\mathcal{C}_+{}^{\mathrm{tr}} & -\mathcal{C}_-{}^{\mathrm{tr}} & \mathcal{A}_- & \mathcal{A}_+ \\ -\mathcal{C}_-{}^{\mathrm{tr}} & -\mathcal{C}_+{}^{\mathrm{tr}} & \mathcal{A}_+ & \mathcal{A}_- \end{pmatrix} \begin{pmatrix} -\Delta(G\circ\widetilde{J})''/2 \\ \phi'' - \overline{(G\circ\widetilde{J})''} \\ \overline{\phi'} - \overline{(G\circ\widetilde{J})'} \\ \Delta\phi'/2 - \Delta(G\circ\widetilde{J})'/2 \end{pmatrix} \\
& -\frac{1}{2}\begin{pmatrix} \overline{\phi'} \\ \Delta\phi'/2 \end{pmatrix}^{\mathrm{tr}} \begin{pmatrix} \Omega_0^2\mathcal{V}(\Omega_0) & 0 \\ 0 & 4\mathcal{V}^{-1}(\Omega_0) \end{pmatrix} \begin{pmatrix} \overline{\phi'} \\ \Delta\phi'/2 \end{pmatrix} + i \begin{pmatrix} \phi'' \\ \overline{\phi'} \\ \Delta\phi'/2 \end{pmatrix}^{\mathrm{tr}} \begin{pmatrix} \Delta\left[\partial_t(mG\circ\widetilde{J}) + \widetilde{y}\right]'' \\ -\Delta\left[\partial_t(mG\circ\widetilde{J}) + \widetilde{y}\right]' \\ -2\overline{\left[\partial_t(mG\circ\widetilde{J}) + \widetilde{y}\right]'} \end{pmatrix} \\
& - iS_{0+\phi\phi}[G_1\circ\widetilde{J}_1,\Phi_1] + iS_{0+\phi\phi}[G_2\circ\widetilde{J}_2,\Phi_2] \Bigg\}. \qquad \text{(IV.6.38)}
\end{aligned}
$$

Defining the matrix

$$
\mathcal{P} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \Omega_0^2\mathcal{V}(\Omega_0) & 0 \\ 0 & 0 & 4\mathcal{V}^{-1}(\Omega_0) \end{pmatrix} \qquad \text{(IV.6.39)}
$$

so that $\mathcal{M}$, defined in (IV.5.33), can be written

$$
\mathcal{M} = \mathcal{P} - i \begin{pmatrix} \mathcal{B}_- & -\mathcal{C}_- & -\mathcal{C}_+ \\ -\mathcal{C}_-{}^{\mathrm{tr}} & \mathcal{A}_- & \mathcal{A}_+ \\ -\mathcal{C}_+{}^{\mathrm{tr}} & \mathcal{A}_+ & \mathcal{A}_- \end{pmatrix}. \qquad \text{(IV.6.40)}
$$

Then (IV.6.38) becomes

$$
\begin{aligned}
Z[J_1,J_2,\Phi_1,\Phi_2] \propto \int & \frac{d\overline{\phi'}d\Delta\phi' d\phi''}{\sqrt{\det(2\pi\mathfrak{C}[\Phi_1])\det(2\pi\mathfrak{C}[\Phi_2])}} \exp\Bigg\{ -\frac{1}{2}\begin{pmatrix} \phi'' \\ \overline{\phi'} \\ \Delta\phi'/2 \end{pmatrix}^{\mathrm{tr}} \mathcal{M} \begin{pmatrix} \phi'' \\ \overline{\phi'} \\ \Delta\phi'/2 \end{pmatrix} \\
& + \begin{pmatrix} \phi'' \\ \overline{\phi'} \\ \Delta\phi'/2 \end{pmatrix}^{\mathrm{tr}} \left[(\mathcal{M}-\mathcal{P})\mathcal{U}[\widetilde{J}] + i\mathcal{W}[\widetilde{J}] + i\begin{pmatrix} \Delta\widetilde{y}'' \\ -\Delta\overline{\widetilde{y}'} \\ -2\overline{\widetilde{y}'} \end{pmatrix} \right] + i\mathcal{Q}[\widetilde{J}_1,\widetilde{J}_2] \Bigg\}, \qquad \text{(IV.6.41)}
\end{aligned}
$$

where

$$\mathcal{U}[\widetilde{J}] = \begin{pmatrix} \overline{(G \circ \widetilde{J})''} \\ \overline{(G \circ \widetilde{J})'} \\ \Delta(G \circ \widetilde{J})'/2 \end{pmatrix} \tag{IV.6.42a}$$

$$\mathcal{W}[\widetilde{J}] = \begin{pmatrix} -\mathcal{B}_+ \Delta(G \circ \widetilde{J})''/2 + \Delta \left[\partial_t (mG \circ \widetilde{J})\right]'' \\ {\mathcal{C}_+}^{\mathrm{tr}} \Delta(G \circ \widetilde{J})''/2 - \Delta \left[\partial_t (mG \circ \widetilde{J})\right]' \\ {\mathcal{C}_-}^{\mathrm{tr}} \Delta(G \circ \widetilde{J})''/2 - 2\overline{\left[\partial_t (mG \circ \widetilde{J})\right]'} \end{pmatrix} \tag{IV.6.42b}$$

$$\begin{aligned} \mathcal{Q}[\widetilde{J}_1, \widetilde{J}_2] = \frac{1}{2} & \begin{pmatrix} \Delta(G \circ \widetilde{J})''/2 \\ \overline{(G \circ \widetilde{J})''} \\ \overline{(G \circ \widetilde{J})'} \\ \Delta(G \circ \widetilde{J})'/2 \end{pmatrix}^{\mathrm{tr}} \begin{pmatrix} \mathcal{B}_- & \mathcal{B}_+ & -\mathcal{C}_+ & -\mathcal{C}_- \\ \mathcal{B}_+ & \mathcal{B}_- & -\mathcal{C}_- & -\mathcal{C}_+ \\ -{\mathcal{C}_+}^{\mathrm{tr}} & -{\mathcal{C}_-}^{\mathrm{tr}} & \mathcal{A}_- & \mathcal{A}_+ \\ -{\mathcal{C}_-}^{\mathrm{tr}} & -{\mathcal{C}_+}^{\mathrm{tr}} & \mathcal{A}_+ & \mathcal{A}_- \end{pmatrix} \begin{pmatrix} \Delta(G \circ \widetilde{J})''/2 \\ \overline{(G \circ \widetilde{J})''} \\ \overline{(G \circ \widetilde{J})'} \\ \Delta(G \circ \widetilde{J})'/2 \end{pmatrix} \\ & - S_{0+\phi\phi}[G_1 \circ \widetilde{J}_1, \Phi_1] + S_{0+\phi\phi}[G_2 \circ \widetilde{J}_2, \Phi_2]. \end{aligned} \tag{IV.6.42c}$$

Completing the square in (IV.6.41) gives

$$\begin{aligned} Z[J_1, J_2, \Phi_1, \Phi_2] = e^{iW_{0+\phi\phi}[\Phi_1,\Phi_2]} \exp \Bigg\{ & \frac{1}{2} \left[\mathcal{U}[\widetilde{J}]^{\mathrm{tr}} (\mathcal{M} - \mathcal{P}) + i\mathcal{W}[\widetilde{J}]^{\mathrm{tr}} + i \begin{pmatrix} \Delta\widetilde{y}'' \\ -\Delta\widetilde{y}' \\ -2\overline{\widetilde{y}'} \end{pmatrix}^{\mathrm{tr}} \right] \\ & \times \mathcal{M}^{-1} \left[(\mathcal{M} - \mathcal{P}) \mathcal{U}[\widetilde{J}] + i\mathcal{W}[\widetilde{J}] + i \begin{pmatrix} \Delta\widetilde{y}'' \\ -\Delta\widetilde{y}' \\ -2\overline{\widetilde{y}'} \end{pmatrix} \right] + i\mathcal{Q}[\widetilde{J}_1, \widetilde{J}_2] \Bigg\} \end{aligned} \tag{IV.6.43}$$

Expanding (IV.6.23b) in a perturbation series, we see that terms beyond the zeroth have at least one factor of λ, from the $\lambda S_{\phi\phi\phi}$. Again, the only way that a perturbative expression could tell us about the non-perturbative result $e^{iW} \ll 1$ is if some of the terms have a λ/β behavior. Thus, we should look for the terms in the exponential of (IV.6.43) which are larger than $\mathcal{O}(1)$ to see if any $\mathcal{O}(\beta^{-1})$ terms can produce significant contributions. The only object which can be larger than $\mathcal{O}(1)$ is $\mathcal{M}^{-1}$ [the matrices $\mathcal{M}$ and $\mathcal{P}$ individually have $\mathcal{V}^{-1}$ eigenvalues, but the combination $\mathcal{M} - \mathcal{P}$ is $\mathcal{O}(1)$]. Since the smallest eigenvalue of $\mathcal{M}$ is $\mathcal{O}(\mathcal{V}) + \mathcal{O}(\lambda)$, $\lambda\mathcal{M}^{-1}$ will also be no larger than $\mathcal{O}(1)$. And since the terms $\widetilde{y}$ and $\widetilde{J} - J = -\widetilde{x} + \dot{\widetilde{y}}$ coming from S_ϕ are $\mathcal{O}(\lambda)$, this means that

$$\begin{aligned} & Z[J_1, J_2, \Phi_1, \Phi_2] = e^{iW_{0+\phi\phi}[\Phi_1,\Phi_2]} \\ & \times \exp\left(\frac{1}{2} \left\{ \mathcal{U}[J]^{\mathrm{tr}} (\mathcal{M} - \mathcal{P}) + i\mathcal{W}[J]^{\mathrm{tr}} \right\} \mathcal{M}^{-1} \left\{ (\mathcal{M} - \mathcal{P}) \mathcal{U}[J] + i\mathcal{W}[J] \right\} + \mathcal{O}(1) \right). \end{aligned} \tag{IV.6.44}$$

Now,

$$\begin{aligned}\left[\mathcal{U}^{\mathrm{tr}}(\mathcal{M}-\mathcal{P})+i\mathcal{W}^{\mathrm{tr}}\right]&\mathcal{M}^{-1}\left[(\mathcal{M}-\mathcal{P})\mathcal{U}+i\mathcal{W}\right]\\ &=\mathcal{U}^{\mathrm{tr}}\mathcal{M}\mathcal{U}-2\mathcal{U}^{\mathrm{tr}}(\mathcal{P}\mathcal{U}-i\mathcal{W})+(\mathcal{U}^{\mathrm{tr}}\mathcal{P}-i\mathcal{W}^{\mathrm{tr}})\mathcal{M}^{-1}(\mathcal{P}\mathcal{U}-i\mathcal{W})\\ &=\mathcal{U}^{\mathrm{tr}}(\mathcal{M}-\mathcal{P})\mathcal{U}-\mathcal{U}^{\mathrm{tr}}\mathcal{P}\mathcal{U}+2i\mathcal{U}^{\mathrm{tr}}\mathcal{W}+(\mathcal{U}^{\mathrm{tr}}\mathcal{P}-i\mathcal{W}^{\mathrm{tr}})\mathcal{M}^{-1}(\mathcal{P}\mathcal{U}-i\mathcal{W})\\ &=-\mathcal{U}^{\mathrm{tr}}\mathcal{P}\mathcal{U}+(\mathcal{U}^{\mathrm{tr}}\mathcal{P}-i\mathcal{W}^{\mathrm{tr}})\mathcal{M}^{-1}(\mathcal{P}\mathcal{U}-i\mathcal{W})+\mathcal{O}(1);\end{aligned} \tag{IV.6.45}$$

if we use (IV.5.36) to write

$$\mathcal{M}^{-1}=\begin{pmatrix}1&0&0\\0&1&0\\-i\alpha^{-1}\mathcal{C}_+{}^{\mathrm{tr}}&i\alpha^{-1}\mathcal{A}_+&1\end{pmatrix}\widetilde{\mathcal{M}}^{-1}\begin{pmatrix}1&0&-i\mathcal{C}_+\alpha^{-1}\\0&1&i\mathcal{A}_+\alpha^{-1}\\0&0&1\end{pmatrix} \tag{IV.6.46}$$

and observe [since $4\mathcal{V}^{-1}(\Omega_0)=\alpha+i\mathcal{A}_-$]

$$\begin{pmatrix}1&0&-i\mathcal{C}_+\alpha^{-1}\\0&1&i\mathcal{A}_+\alpha^{-1}\\0&0&1\end{pmatrix}\mathcal{P}=\begin{pmatrix}0&0&-i\mathcal{C}_+(1+i\alpha^{-1}\mathcal{A}_-)\\0&\Omega_0^2\mathcal{V}(\Omega_0)&i\mathcal{A}_+(1+i\alpha^{-1}\mathcal{A}_-)\\0&0&4\mathcal{V}^{-1}(\Omega_0)\end{pmatrix}, \tag{IV.6.47}$$

we have

$$\begin{aligned}&-\mathcal{U}^{\mathrm{tr}}\mathcal{P}\mathcal{U}+(\mathcal{U}^{\mathrm{tr}}\mathcal{P}-i\mathcal{W}^{\mathrm{tr}})\mathcal{M}^{-1}(\mathcal{P}\mathcal{U}-i\mathcal{W})=-\mathcal{U}^{\mathrm{tr}}\begin{pmatrix}0&0&0\\0&\Omega_0^2\mathcal{V}(\Omega_0)&0\\0&0&4\mathcal{V}^{-1}(\Omega_0)\end{pmatrix}\mathcal{U}\\ &+\left[\mathcal{U}^{\mathrm{tr}}\begin{pmatrix}0&0&0\\0&\Omega_0^2\mathcal{V}(\Omega_0)&0\\-i(1+i\mathcal{A}_-\alpha^{-1})\mathcal{C}_+^{\mathrm{tr}}&i(1+i\mathcal{A}_-\alpha^{-1})\mathcal{A}_+&4\mathcal{V}^{-1}(\Omega_0)\end{pmatrix}\right.\\ &\qquad\qquad\left.-i\mathcal{W}^{\mathrm{tr}}\begin{pmatrix}1&0&0\\0&1&0\\-i\alpha^{-1}\mathcal{C}_+{}^{\mathrm{tr}}&i\alpha^{-1}\mathcal{A}_+&1\end{pmatrix}\right]\\ &\times\widetilde{M}^{-1}\left[\begin{pmatrix}0&0&-i\mathcal{C}_+(1+i\alpha^{-1}\mathcal{A}_-)\\0&\Omega_0^2\mathcal{V}(\Omega_0)&i\mathcal{A}_+(1+i\alpha^{-1}\mathcal{A}_-)\\0&0&4\mathcal{V}^{-1}(\Omega_0)\end{pmatrix}\mathcal{U}-i\begin{pmatrix}1&0&-i\mathcal{C}_+\alpha^{-1}\\0&1&i\mathcal{A}_+\alpha^{-1}\\0&0&1\end{pmatrix}\mathcal{W}\right].\end{aligned} \tag{IV.6.48}$$

Because the matrices $\mathcal{P}$ and $\aleph_0-i\lambda\aleph_1[\Delta\Phi])^{-1}\oplus\alpha^{-1}$ are in block diagonal form, we split up the expression (IV.6.48) into

$$\begin{aligned}&-\mathcal{U}^{\mathrm{tr}}\begin{pmatrix}0&0&0\\0&\Omega_0^2\mathcal{V}(\Omega_0)&0\\0&0&0\end{pmatrix}\mathcal{U}\\ &+\left[\mathcal{U}^{\mathrm{tr}}\begin{pmatrix}0&0\\0&\Omega_0^2\mathcal{V}(\Omega_0)\\-i(1+i\mathcal{A}_-\alpha^{-1})\mathcal{C}_+^{\mathrm{tr}}&i(1+i\mathcal{A}_-\alpha^{-1})\mathcal{A}_+\end{pmatrix}-i\mathcal{W}^{\mathrm{tr}}\begin{pmatrix}1&0\\0&1\\-i\alpha^{-1}\mathcal{C}_+{}^{\mathrm{tr}}&i\alpha^{-1}\mathcal{A}_+\end{pmatrix}\right]\\ &\times(\aleph_0-i\lambda\aleph_1[\Delta\Phi])^{-1}\left[\begin{pmatrix}0&0&-i\mathcal{C}_+(1+i\alpha^{-1}\mathcal{A}_-)\\0&\Omega_0^2\mathcal{V}(\Omega_0)&i\mathcal{A}_+(1+i\alpha^{-1}\mathcal{A}_-)\end{pmatrix}\mathcal{U}-i\begin{pmatrix}1&0&-i\mathcal{C}_+\alpha^{-1}\\0&1&i\mathcal{A}_+\alpha^{-1}\end{pmatrix}\mathcal{W}\right].\end{aligned} \tag{IV.6.49a}$$

and

$$-\mathcal{U}_3^{\rm tr} 4\mathcal{V}^{-1}(\Omega_0)\mathcal{U}_3 + \left[\mathcal{U}_3^{\rm tr} 4\mathcal{V}^{-1}(\Omega_0) - i\mathcal{W}_3^{\rm tr}\right]\alpha^{-1}\left[4\mathcal{V}^{-1}(\Omega_0)\mathcal{U}_3 - i\mathcal{W}_3\right], \tag{IV.6.49b}$$

where $\mathcal{U}_3$ is the bottom third of $\mathcal{U}$,

$$\mathcal{U}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}^{\rm tr} \mathcal{U}. \tag{IV.6.50}$$

Using $4\mathcal{V}^{-1}(\Omega_0) = \alpha + i\mathcal{A}_-$, (IV.6.49b) can be converted, noting

$$\begin{aligned} -4\mathcal{V}^{-1}(\Omega_0) + 4\mathcal{V}^{-1}(\Omega_0)\alpha^{-1}4\mathcal{V}^{-1}(\Omega_0) =& -\alpha - i\mathcal{A}_- + \alpha + 2i\mathcal{A}_- - \mathcal{A}_-\alpha^{-1}\mathcal{A}_- \\ =& i\mathcal{A}_- - \mathcal{A}_-\alpha^{-1}\mathcal{A}_-, \end{aligned} \tag{IV.6.51}$$

to

$$\mathcal{U}_3^{\rm tr}(i\mathcal{A}_- - \mathcal{A}_-\alpha^{-1}\mathcal{A}_-)\mathcal{U}_3 - 2i\mathcal{W}_3^{\rm tr}(1 + i\alpha^{-1}\mathcal{A}_-)\mathcal{U}_3 - \mathcal{W}_3^{\rm tr}\alpha^{-1}\mathcal{W}_3 = \mathcal{O}(1), \tag{IV.6.52}$$

This leaves us with (IV.6.49a), making the exponential in (IV.6.43)

$$\begin{aligned} &\frac{1}{2}\left[\mathcal{U}^{\rm tr}\begin{pmatrix} 0 & 0 \\ 0 & \Omega_0^2\mathcal{V}(\Omega_0) \\ -i\mathcal{C}_+^{\rm tr} & i\mathcal{A}_+ \end{pmatrix} - i\mathcal{W}^{\rm tr}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -i\alpha^{-1}{\mathcal{C}_+}^{\rm tr} & i\alpha^{-1}\mathcal{A}_+ \end{pmatrix}\right](\aleph_0 - i\lambda\aleph_1[\Delta\Phi])^{-1} \\ &\qquad\times\left[\begin{pmatrix} 0 & 0 & -i\mathcal{C}_+ \\ 0 & \Omega_0^2\mathcal{V}(\Omega_0) & i\mathcal{A}_+ \end{pmatrix}\mathcal{U} - i\begin{pmatrix} 1 & 0 & -i\mathcal{C}_+\alpha^{-1} \\ 0 & 1 & i\mathcal{A}_+\alpha^{-1} \end{pmatrix}\mathcal{W}\right] + \mathcal{O}(1); \end{aligned} \tag{IV.6.53}$$

inserting (IV.6.42a) and (IV.6.42b), we find

$$\begin{aligned} &\begin{pmatrix} 0 & 0 & -i\mathcal{C}_+ \\ 0 & \Omega_0^2\mathcal{V}(\Omega_0) & i\mathcal{A}_+ \end{pmatrix}\mathcal{U}[J] - i\begin{pmatrix} 1 & 0 & -i\mathcal{C}_+\alpha^{-1} \\ 0 & 1 & i\mathcal{A}_+\alpha^{-1} \end{pmatrix}\mathcal{W}[J] \\ &= \begin{pmatrix} -i\mathcal{C}_+\Delta(G\circ J)'/2 + i\mathcal{B}_+\Delta(G\circ J)''/2 - i\Delta\left[\partial_t(mG\circ J)\right]'' \\ i\mathcal{A}_+\Delta(G\circ J)'/2 - i{\mathcal{C}_+}^{\rm tr}\Delta(G\circ J)''/2 + i\Delta\left[\partial_t(mG\circ J)\right]' \end{pmatrix} \\ &\qquad + \mathcal{V}(\Omega_0)\begin{pmatrix} \mathcal{V}^{-1}(\Omega_0)\mathcal{C}_+\alpha^{-1}\left\{{\mathcal{C}_-}^{\rm tr}\Delta(G\circ J)''/2 - 2\overline{[\partial_t(mG\circ J)]'}\right\} \\ \Omega_0^2\overline{(G\circ J)'} - \mathcal{V}^{-1}(\Omega_0)\mathcal{A}_+\alpha^{-1}\left\{{\mathcal{C}_-}^{\rm tr}\Delta(G\circ J)''/2 - 2\overline{[\partial_t(mG\circ J)]'}\right\} \end{pmatrix}, \end{aligned} \tag{IV.6.54}$$

so again discarding $\mathcal{O}(1)$ terms {including $(\aleph_0 - i\lambda\aleph_1[\Delta\Phi])^{-1}\mathcal{V}(\Omega_0)$}, we end up with

$$\begin{aligned} &Z[J_1, J_2, \Phi_1, \Phi_2] \\ &\quad= e^{iW_{0+\phi\phi}[\Phi_1,\Phi_2]}\exp\left[-\frac{1}{2}\begin{pmatrix} -C_0G_0\circ(\Delta J)' + B_0G_0\circ(\Delta J)'' - \dot{G}_0\circ(\Delta J)'' \\ A_0G_0\circ(\Delta J)' - C_0G_0\circ(\Delta J)'' + \dot{G}_0\circ(\Delta J)' \end{pmatrix}^{\rm tr}\right. \\ &\qquad\left.\times(\aleph_0 - i\lambda\aleph_1[\Delta\Phi])^{-1}\begin{pmatrix} -C_0G_0\circ(\Delta J)' + B_0G_0\circ(\Delta J)'' - \dot{G}_0\circ(\Delta J)'' \\ A_0G_0\circ(\Delta J)' - C_0G_0\circ(\Delta J)'' + \dot{G}_0\circ(\Delta J)' \end{pmatrix} + \mathcal{O}(1)\right]. \end{aligned} \tag{IV.6.55}$$

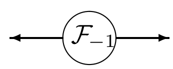
The fact that the leading term in the exponential in (IV.6.55) depends only upon $\Delta J = J_1 - J_2$ is crucial, because of the operator

$$\lambda S_{\phi\phi\phi}\left[\frac{1}{i}\frac{\mathcal{D}}{\mathcal{D}J_1}, \Phi_1\right] - \lambda S_{\phi\phi\phi}\left[-\frac{1}{i}\frac{\mathcal{D}}{\mathcal{D}J_2}, \Phi_2\right] \tag{IV.6.56}$$

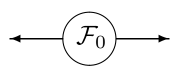
in (IV.6.23b), which annihilates any functional depending only on the combination $J_1 - J_2$. If *all* of the terms in the exponential in Z were functions of ΔJ alone, that would mean that $e^{iW} = e^{iW_3}$; however, there are $\mathcal{O}(1)$ terms in the exponential which depend on $\overline{J}$. The situation can be written as

$$Z[J_1, J_2, \Phi_1, \Phi_2] = e^{iW_{0+\phi\phi}[\Phi_1,\Phi_2]} \exp\left(\frac{1}{2}\Delta J \circ \mathcal{F}_{-1} \circ \Delta J + \frac{1}{2} J \circ \mathcal{F}_0 \circ J + J \circ \mathcal{G}_0\right), \tag{IV.6.57}$$

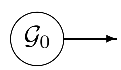
where $\frac{1}{2}\Delta J \circ \mathcal{F}_{-1} \circ \Delta J$ is the argument of the exponential in (IV.6.55), and $\frac{1}{2} J \circ \mathcal{F}_0 \circ J$ and $J \circ \mathcal{G}_0$ are the quadratic and linear terms of $\mathcal{O}(1)$. Thinking in terms of a diagrammatic expansion, this means that there are three kinds of "propagators" in Z:



(IV.6.58a)

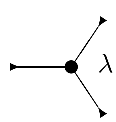


(IV.6.58b)



(IV.6.58c)

(note that the last is not truly a propagator, since it accepts only one "input"). These are used to connect the vertices, which all have the form

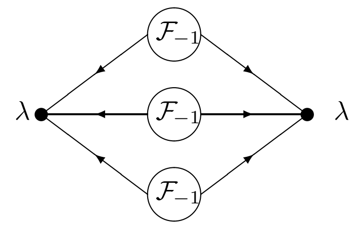


(IV.6.59)

A term in the series which has more λ vertices than $\mathcal{F}_{-1}$ propagators will be perturbatively small, one with the same number will be $\mathcal{O}(1)$, and one with more $\mathcal{F}_{-1}$ propagators than λ vertices will be able to disrupt the perturbative analysis and have an impact upon e^{iW}. We can make a list of the objects in the theory by their order in perturbation theory and number of legs (with the legs on propagators counted negative so that a closed diagram has zero net legs):

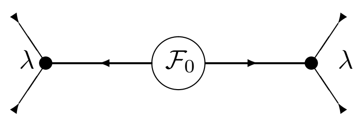
Graph	Order	Legs
(IV.6.58a)	-1	-2
(IV.6.58b)	0	-2
(IV.6.58c)	0	-1
(IV.6.59)	1	3

Since the vertex (IV.6.59) has three legs and the propagator (IV.6.58a) has minus two, we'd expect divergent graphs starting with

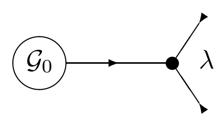


(IV.6.60)

However, in this case we have just the situation described above: all of the propagators depend only on ΔJ, so the graph vanishes. This sort of identity places the restriction that at least one leg of a vertex must be coupled to an $\mathcal{F}_0$ or $\mathcal{G}_0$ propagator. This means that we must abandon (IV.6.59) by itself and use as our primitive vertices



(IV.6.61a)



(IV.6.61b)

which makes the pieces out of which non-vanishing graphs can be constructed

Graph	Order	Legs
(IV.6.58a)	-1	-2
(IV.6.58b)	0	-2
(IV.6.58c)	0	-1
(IV.6.61a)	2	4
(IV.6.61b)	1	2

Now the most divergent graph which can be constructed with zero net legs is $\mathcal{O}(1)$.

This means that, perturbatively, the influence functional is

$$e^{iW[\Phi_1,\Phi_2]} = \mathcal{O}(1) \times e^{iW_3[\Phi_1,\Phi_2]}, \tag{IV.6.62}$$

so, perturbatively at least,

$$\left| e^{iW[\Phi_1,\Phi_2]} \right| \lesssim \left\{ 1 + \mathrm{Tr}\left(\lambda \aleph_0^{-1} \aleph_1[\Delta\Phi] \right)^2 \right\}^{-1/4}. \tag{IV.6.63}$$

A word about the perturbative analysis

The conclusion that

$$e^{iW[\Phi_1,\Phi_2]-iW_3[\Phi_1,\Phi_2]} = \mathcal{O}(1) \tag{IV.6.62}$$

is based upon an upper limit on each term in the perturbation series (the first term is obviously unity). There are two ways this analysis could fail. First, there may be cancellation among the various $\mathcal{O}(1)$ terms causing the net expression to be a higher order in λ or β. Since this would only make $\left|e^{iW}\right|$ *smaller* than our estimate, it would only improve the upper limit given by (IV.6.63)

The second is more problematic. While each individual term is at most $\mathcal{O}(1)$, the entire infinite series could be quite large, counteracting the tendency of $e^{iW_{0+\phi\phi}}$ to become small. This is a shortcoming of the perturbative analysis, and there's not a lot to be done, other than to tackle the non-perturbative problem.[6] Note, however, that we can say with confidence that $\left|e^{iW-iW_3}\right|$ does not have terms which are $\mathcal{O}(\lambda^2/\beta^2)$, which could directly cancel similar terms in the expansion. So if $\left|e^{iW-iW_3}\right|$ becomes large, it is not in the same way which $\left|e^{iW_3}\right| = \left|e^{iW_{0+\phi\phi}}\right|$ becomes small.

IV.7 Interpretation

IV.7.1 Which modes are suppressed?

Having determined that the influence functional is bounded from above by

$$\left|e^{iW[\Phi_1,\Phi_2]}\right| \lesssim \left\{1 + \mathrm{Tr}\left(\lambda\aleph_0^{-1}\aleph_1[\Delta\Phi]\right)^2\right\}^{-1/4}, \tag{IV.6.63}$$

and hence becomes small when

$$\begin{aligned}
\mathrm{Tr}(\lambda\aleph_0^{-1}\aleph_1)^2 = &\int_0^\infty dq \iint q^2 d^2\Omega_{\hat{q}} \int_{-q}^{q} dk_- \int_{2k_c+|k_-|}^{\infty} dk_+ \frac{\coth\beta\frac{k_++k_-}{4}\coth\beta\frac{k_+-k_-}{4}}{(2\pi)^2 512 q} \\
&\times\Bigg\{\left|(q^2-k_-^2)\int_{-T/2}^{T/2} dt\ell\Delta\varphi_{\mathbf{q}}(t)e^{ik_-t} - i4k_-\left[e^{i2k_-t}\ell\Delta\varphi_{\mathbf{q}}(t)\right]_{-T/2}^{T/2}\right|^2 \\
&+\left|(q^2-k_-^2)\int_{-T/2}^{T/2} dt\ell\Delta\varphi_{\mathbf{q}}(t)e^{-ik_-t} + i4k_-\left[e^{-i2k_-t}\ell\Delta\varphi_{\mathbf{q}}(t)\right]_{-T/2}^{T/2}\right|^2 \\
&+\left|(k_+^2-q^2)\int_{-T/2}^{T/2} dt\ell\Delta\varphi_{\mathbf{q}}(t)e^{ik_+t} + i4k_+\left[e^{i2k_+t}\ell\Delta\varphi_{\mathbf{q}}(t)\right]_{-T/2}^{T/2}\right|^2 \\
&+\left|(k_+^2-q^2)\int_{-T/2}^{T/2} dt\ell\Delta\varphi_{\mathbf{q}}(t)e^{-ik_+t} - i4k_+\left[e^{-i2k_+t}\ell\Delta\varphi_{\mathbf{q}}(t)\right]_{-T/2}^{T/2}\right|^2\Bigg\}.
\end{aligned} \tag{IV.5.74}$$

becomes large, we would like to consider when that happens. Looking at (IV.5.74), and disregarding the surface terms, we see that not all of the space/time modes

$$\Delta\varphi_{\mathbf{q}\omega} = \int_{-T/2}^{T/2} \frac{dt}{\sqrt{2\pi}} \Delta\varphi_{\mathbf{q}}(t)e^{i\omega t} \tag{IV.7.1}$$

[6] For instance, we can't use (IV.5.37) to conclude that the $\mathcal{O}(1)$ factor in (IV.6.62) is unity, since that would involve an illegal interchange of the $\beta \to 0$ and $\lambda \to 0$ limits.

appear. The first two terms include only modes where $|\omega| = |k_-| \leq q$, while the last two are limited to modes where $|\omega| = |k_+| \geq 2k_c$. This is illustrated in Fig. IV.3. Just as our coarse graining considers only long-wavelength modes ($q \leq k_c$), it is reasonable to focus on long-period modes ($|\omega| \leq k_c$) as well. Thus the limit of interest comes from the first two terms, and we write

$$\begin{aligned}
\mathrm{Tr}(\lambda\aleph_0^{-1}\aleph_1)^2 \geq & \int_0^\infty dq \iint q^2 d^2\Omega_{\hat{q}} \int_{-q}^{q} dk_- \int_{2k_c+|k_-|}^{\infty} dk_+ \frac{\coth\beta\frac{k_++k_-}{4}\coth\beta\frac{k_+-k_-}{4}}{(2\pi)^2 512 q} \\
& \times\Bigg\{ \left|(q^2-k_-^2)\sqrt{2\pi}\ell\Delta\varphi_{\mathbf{q}k_-} - i4k_-\left[e^{i2k_-t}\ell\Delta\varphi_{\mathbf{q}}(t)\right]_{-T/2}^{T/2}\right|^2 \\
& \quad + \left|(q^2-k_-^2)\sqrt{2\pi}\ell\Delta\varphi_{\mathbf{q},-k_-} + i4k_-\left[e^{-i2k_-t}\ell\Delta\varphi_{\mathbf{q}}(t)\right]_{-T/2}^{T/2}\right|^2\Bigg\}. \\
= & \int_0^\infty dq \iint q^2 d^2\Omega_{\hat{q}} \int_{-q}^{q} dk_- \left|(q^2-k_-^2)\sqrt{2\pi}\ell\Delta\varphi_{\mathbf{q}k_-} - i4k_-\left[e^{i2k_-t}\ell\Delta\varphi_{\mathbf{q}}(t)\right]_{-T/2}^{T/2}\right|^2 \\
& \times \int_{2k_c+|k_-|}^{\infty} dk_+ \frac{\coth\beta\frac{k_++k_-}{4}\coth\beta\frac{k_+-k_-}{4}}{(2\pi)^2 256 q}.
\end{aligned} \tag{IV.7.2}$$

The factor

$$R = \int_{2k_c+|k_-|}^{\infty} dk_+ \frac{\coth\frac{\beta k_1}{2}\coth\frac{\beta k_2}{2}}{(2\pi)^2 256 q} \tag{IV.7.3}$$

can be evaluated, to leading order in β, by noting that

$$\coth\eta_1\coth\eta_2 = \frac{\cosh\eta_1\cosh\eta_2}{\sinh\eta_1\sinh\eta_2} = \frac{\cosh\eta_+ + \cosh\eta_-}{\cosh\eta_+ - \cosh\eta_-} = 1 + \frac{2\cosh\eta_-}{\cosh\eta_+ - \cosh\eta_-}, \tag{IV.7.4}$$

so that

$$R = \int_{2k_c+|k_-|}^{\infty} \frac{dk_+}{(2\pi)^2 256 q}\left(1 + \frac{2\cosh\frac{\beta k_-}{2}}{\cosh\frac{\beta k_+}{2} - \cosh\frac{\beta k_-}{2}}\right) = R_0 + \int_{2k_c+|k_-|}^{\infty} \frac{dk_+}{(2\pi)^2 256 q}. \tag{IV.7.5}$$

Now,

$$\begin{aligned}
R_0 = & \int_{2k_c+|k_-|}^{\infty} \frac{dk_+}{(2\pi)^2 256 q}\frac{2\cosh\frac{\beta k_-}{2}}{\cosh\frac{\beta k_+}{2} - \cosh\frac{\beta k_-}{2}} \\
= & \frac{4\cosh\frac{\beta k_-}{2}}{(2\pi)^2 256 q\beta\sinh\frac{\beta|k_-|}{2}} \ln\left(\frac{\sinh\beta\frac{k_+-|k_-|}{4}}{\sinh\beta\frac{k_++|k_-|}{4}}\right)\Bigg|_{2k_c+|k_-|}^{\infty} \\
= & \frac{4\coth\frac{\beta|k_-|}{2}}{(2\pi)^2 256 q\beta} \ln\left(\frac{\sinh\beta\frac{k_c+|k_-|}{2}}{\sinh\frac{\beta k_c}{2}e^{\beta|k_-|/2}}\right).
\end{aligned} \tag{IV.7.6}$$

Again, since we only expect a useful answer when small β causes perturbation theory to break down, we look at the leading terms in β, working in the high-temperature limit $\beta k_c \gg 1$.

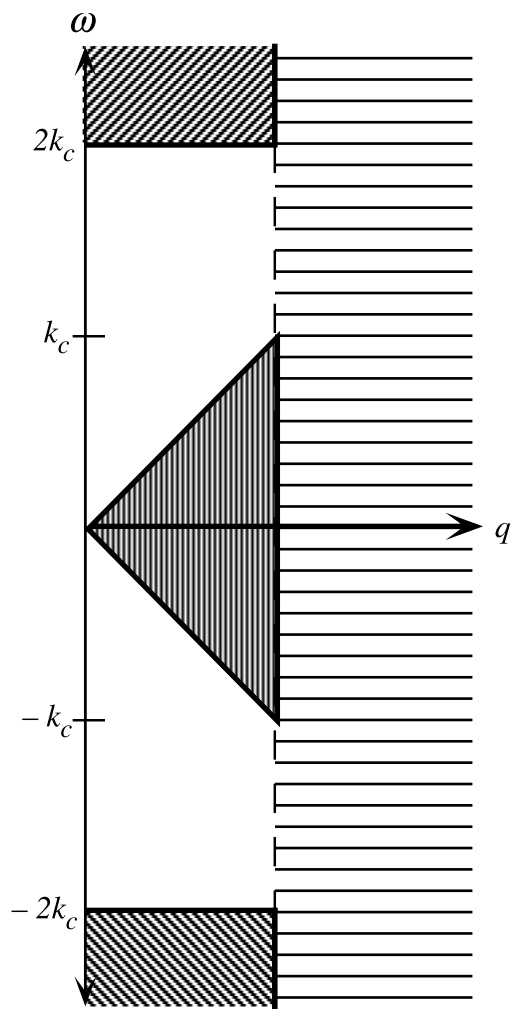


Figure IV.3: The modes represented in (IV.5.74), plotted by their ω and q values. The modes with $q \geq k_c$ are traced over, and so that region is shaded horizontally. The first two terms in (IV.5.74) can suppress modes with $|\omega| \leq q$, which are shaded vertically, the third can suppress modes which have $\omega \geq k_c$ and the fourth, $\omega \leq -k_c$; these last two are shaded diagonally. Since we are concerned with coarse grainings of low temporal frequency ω as well as spatial frequency q, the first two terms are the ones of interest.

(See Sec. IV.7.2 for the physical significance of this.) In this limit, (IV.7.6) becomes

$$R_0 = \frac{8}{(2\pi)^2 256 q\beta^2 |k_-|} \ln\left(1 + \frac{|k_-|}{k_c}\right); \tag{IV.7.7}$$

since $R - R_0$ is independent of β, the leading term in R is[7]

$$R = \frac{1}{(2\pi)^2 32 q\beta^2 |k_-|} \ln\left(1 + \frac{|k_-|}{k_c}\right), \tag{IV.7.8}$$

so

$$\begin{aligned} \mathrm{Tr}(\lambda\aleph_0^{-1}\aleph_1)^2 \gtrsim \int_0^\infty dq \iint q^2 d^2\Omega_{\hat{q}} \int_{-q}^{q} \frac{d\omega}{(2\pi)^2 32 q\beta^2 |\omega|} \ln\left(1 + \frac{|\omega|}{k_c}\right) \\ \times \left|(q^2-\omega^2)\sqrt{2\pi}\ell\Delta\varphi_{\mathbf{q}\omega} - i4\omega\left[e^{i2\omega t}\ell\Delta\varphi_{\mathbf{q}}(t)\right]_{-T/2}^{T/2}\right|^2. \end{aligned} \tag{IV.7.9}$$

IV.7.2 Practical coarse grainings

The physical scales

The expression (IV.7.9) has three parameters, k_c, β and T which are not integrated over. The scale k_c for division into SWMs and LWMs can be tailored to the coarse graining to give the strongest possible results, while the other two are features of the model. As alluded to in Sec. IV.3.3, the time scale T over which we expect the Minkowski space model to be valid should be slightly below the Hubble scale H_0^{-1}. In suitable units, this gives

$$T \lesssim H_0^{-1} \sim 10^{10}\,\mathrm{yr} \sim 10^{29}\,\mathrm{cm}. \tag{IV.7.10}$$

This is so large that it allows us to set T much larger than all the other scales in the problem. In particular, it means that the cross terms in

$$\begin{aligned} &\left|(q^2-\omega^2)\sqrt{2\pi}\ell\Delta\varphi_{\mathbf{q}\omega} - i4\omega\left[e^{i2\omega t}\ell\Delta\varphi_{\mathbf{q}}(t)\right]_{-T/2}^{T/2}\right|^2 \\ &\quad = 2\pi\left|(q^2-\omega^2)\ell\Delta\varphi_{\mathbf{q}\omega}\right|^2 + 16\omega^2\left|e^{i\omega T}\ell\Delta\varphi''_{\mathbf{q}} - e^{-i\omega T}\ell\Delta\varphi'_{\mathbf{q}}\right|^2 \\ &\qquad + i\sqrt{2\pi}(q^2-\omega^2)\ell\Delta\varphi_{\mathbf{q}\omega}4\omega\left(e^{-i\omega T}\ell\Delta\varphi''^{*}_{\mathbf{q}} - e^{i\omega T}\ell\Delta\varphi'^{*}_{\mathbf{q}}\right) \\ &\qquad - i\sqrt{2\pi}(q^2-\omega^2)\ell\Delta\varphi^{*}_{\mathbf{q}\omega}4\omega\left(e^{i\omega T}\ell\Delta\varphi''_{\mathbf{q}} - e^{-i\omega T}\ell\Delta\varphi'_{\mathbf{q}}\right) \end{aligned} \tag{IV.7.11}$$

will oscillate rapidly and vanish when ω is integrated over, leaving

$$= 2\pi\left|(q^2-\omega^2)\ell\Delta\varphi_{\mathbf{q}\omega}\right|^2 + 16\omega^2\left|e^{i\omega T}\ell\Delta\varphi''_{\mathbf{q}} - e^{-i\omega T}\ell\Delta\varphi'_{\mathbf{q}}\right|^2 \geq 2\pi\left|(q^2-\omega^2)\ell\Delta\varphi_{\mathbf{q}\omega}\right|. \tag{IV.7.12}$$

[7] Of course, this is a dubious approximation, since $R - R_0$, while down by a factor of β^2 from R_0, is ultraviolet divergent. However, any suitable well-behaved regulation of the result will give a result which agrees with R_0 to $\mathcal{O}(\beta^{-2})$ when the $\beta \to 0$ limit is taken before the cutoff limit. Note also that our perturbative analysis has ignored terms like $\lambda^2(R - R_0)$, which are perturbatively small in λ without having corresponding factors of β. One might hope that such terms will cancel the divergence in $R - R_0$. However, this turns out not to be the case, as can be seen by calculating all of the $\mathcal{O}(\lambda^2)$ terms in $e^{iW_{0+\phi\phi}}$.

Turning our attention to the inverse temperature β, we might reasonably treat the high-temperature thermal state ρ_ϕ as corresponding to the cosmic graviton background radiation[50], which has a temperature on the order of $1\,\mathrm{K}$. This means that in suitable units,

$$\beta \sim \frac{1}{1\,\mathrm{K}} \sim \frac{1}{10^{-4}\,\mathrm{eV}} \sim 10^{-1}\,\mathrm{cm}. \tag{IV.7.13}$$

This is the most severe limit to the usefulness of the calculations in this chapter. It means that to be in the high-temperature limit $\beta k_c \ll 1$, we need to have the cutoff scale k_c^{-1} dividing "short" and "long" wavelengths be above the millimeter scale. While we don't expect to have laboratory data on millimeter-scale oscillations of vacuum gravity any time soon (contrast this scale to the length corresponding to a typical component of the curvature tensor at the surface of a $1M_\odot$ black hole, which is $GM_\odot \sim 1\,\mathrm{km} \sim 10^6\beta$), it might be a bit surprising to learn that coarse grainings corresponding to micron-scale variations in the gravitational field do not decohere. At any rate, that is *not* the prediction of this chapter, even assuming that the results for the toy model are an accurate indicator of the behavior of the actual theory. First, this analysis only applies to decoherence of the vacuum gravitational field induced by gravity itself. If the gravitational field is coupled to some form of matter, unobserved modes of the matter can also induce decoherence, as described in [27] and [26]. In addition, our perturbative analysis of vacuum gravity simply cannot make fruitful predictions outside of the perturbative regime. It is quite possible that for lower temperatures, non-perturbative effects can cause the influence functional to become small for large $\Delta\varphi$.

Field averages

A simple quantity by which we'd like to coarse grain would be a field average

$$\langle \ell\varphi \rangle = \int d^3x \int_{-T/2}^{T/2} dt f(\mathbf{x},t)\ell\varphi(\mathbf{x},t), \tag{IV.7.14}$$

where $f(\mathbf{x},t)$ is some sort of smoothing function with spatial width Δx and temporal width Δt obeying

$$\int d^3x \int_{-T/2}^{T/2} dt f(\mathbf{x},t) = 1. \tag{IV.7.15}$$

In terms of Fourier modes, this becomes

$$\langle \ell\varphi \rangle = \int d^3q\, d\omega\, f^*_{\mathbf{q}\omega}\ell\varphi_{\mathbf{q}\omega}, \tag{IV.7.16}$$

where I have approximated the sum over ω values separated by $\delta\omega = 2\pi/T$ by an integral, and assumed that $f(\mathbf{x},t)$ vanishes as $t \to \pm T/2$, so that it is acceptable to replace the field $\varpi_{\mathbf{q}}(t)$ by its periodic counterpart

$$\varphi^{\mathrm{P}}_{\mathbf{q}}(t) = \int d\omega\varphi_{\mathbf{q}\omega}e^{-i\omega t} = \begin{cases} \frac{\varphi''_{\mathbf{q}}+\varphi'_{\mathbf{q}}}{2} & t = \pm\frac{T}{2} \\ \varpi_{\mathbf{q}}(t) & -\frac{T}{2} < t < \frac{T}{2}. \end{cases} \tag{IV.7.17}$$

The normalization condition (IV.7.15) becomes $f_{\mathbf{00}} = (2\pi)^{-2}$, so a useful field average might be

$$\langle \ell\varphi \rangle = \int_0^{\Delta q/2} dq \iint q^2 d^2\Omega_{\hat{q}} \int_{-\Delta\omega/2}^{\Delta\omega/2} d\omega \varphi_{\mathbf{q}\omega}, \tag{IV.7.18}$$

where the width of the smoothing function in Fourier space is

$$\Delta q \sim \frac{1}{2\Delta x} \tag{IV.7.19a}$$

$$\Delta\omega \sim \frac{1}{2\Delta t} \tag{IV.7.19b}$$

and the origin of the spatial coördinates has been chosen to correspond with the center of $f(\mathbf{x}, t)$. If we shift the center of the group of Fourier modes by $\mathbf{q}_0$ and ω_0, while keeping the mode volume the same, we get another dimensionless quantity

$$\widetilde{\langle \ell\varphi \rangle} = \int_{q_0-\Delta q/2}^{q_0+\Delta q/2} dq \iint\limits_{\Omega} q^2 d^2\Omega_{\hat{q}} \int_{\omega_0-\Delta\omega/2}^{\omega_0+\Delta\omega/2} d\omega \frac{\varphi_{\mathbf{q}\omega}}{(2\pi)^2}, \tag{IV.7.20}$$

where the solid angle integrated over is centered about $\hat{q}_0$ and is chosen to preserve the mode volume:

$$\frac{4\pi(\Delta q/2)^3}{3} = \Omega \int_{q_0-\Delta q/2}^{q_0+\Delta q/2} q^2 dq = \Omega \frac{(q_0 + \Delta q/2)^3 - (q_0 - \Delta q/2)^3}{3} \tag{IV.7.21}$$

so

$$\Omega = \frac{\pi(\Delta q)^2}{6q_0^2 + (\Delta q)^2/2}. \tag{IV.7.22}$$

The influence phase

Now we can cast (IV.7.9) into a useful form, so long as $q_0 - \Delta q/2 \geq |\omega_0| + \Delta\omega$:

$$\begin{aligned} \mathrm{Tr}(\lambda\aleph_0^{-1}\aleph_1)^2 &\gtrsim \int_0^{\infty} dq \iint q^2 d^2\Omega_{\hat{q}} \int_{-q}^{q} \frac{d\omega}{(2\pi)^2 32 q\beta^2 |\omega|} \ln\left(1 + \frac{|\omega|}{k_c}\right) \left|(q^2 - \omega^2)\sqrt{2\pi}\ell\Delta\varphi_{\mathbf{q}\omega}\right|^2 \\ &\geq \int_{q_0-\Delta q/2}^{q_0+\Delta q/2} dq \iint\limits_{\Omega} q^2 d^2\Omega_{\hat{q}} \int_{\omega_0-\Delta\omega/2}^{\omega_0+\Delta\omega/2} d\omega \frac{(q^2-\omega^2)^2}{(2\pi)32q\beta^2 |\omega|} \ln\left(1 + \frac{|\omega|}{k_c}\right) |\ell\Delta\varphi_{\mathbf{q}\omega}|^2 . \end{aligned} \tag{IV.7.23}$$

The strongest result will be obtained if we take $k_c = q_0 + \Delta q/2$. If $\Delta\omega$ and Δq are small relative to ω_0 and q_0 (which means large Δt and Δx), we can approximate

$$\begin{aligned}\mathrm{Tr}(\lambda\aleph_0^{-1}\aleph_1)^2 \gtrsim&\Theta(q_0-|\omega_0|)\frac{(q_0^2-\omega_0^2)^2}{(2\pi)32q_0\beta^2\,|\omega_0|}\ln\left(1+\frac{|\omega_0|}{q_0}\right)\\ &\times\int_{q_0-\Delta q/2}^{q_0+\Delta q/2}dq\iint_{\Omega}q^2d^2\Omega_{\hat{q}}\int_{\omega_0-\Delta\omega/2}^{\omega_0+\Delta\omega/2}d\omega\,\left|\ell\Delta\varphi_{\mathbf{q}\omega}\right|^2\\ \approx&\Theta(q_0-|\omega_0|)\frac{(q_0^2-\omega_0^2)^2}{(2\pi)32q_0\beta^2\,|\omega_0|}\ln\left(1+\frac{|\omega_0|}{q_0}\right)\frac{(2\pi)^2\left|\widetilde{\langle\ell\Delta\varphi\rangle}\right|^2}{\pi\Delta\omega(\Delta q)^3/6}\end{aligned} \tag{IV.7.24}$$

so that the influence phase is bounded by

$$\left|e^{iW[\Phi_1,\Phi_2]}\right|\lesssim\left\{1+\frac{3\pi(q_0^2-\omega_0^2)^2\left|\widetilde{\langle\ell\Delta\varphi\rangle}\right|^2}{8q_0\beta^2\,|\omega_0|\,\Delta\omega(\Delta q)^3}\ln\left(1+\frac{|\omega_0|}{q_0}\right)\right\}^{-1/4}. \tag{IV.7.25}$$

This means that if

$$\frac{3\pi(q_0^2-\omega_0^2)^2\left|\widetilde{\langle\ell\Delta\varphi\rangle}\right|^2}{8q_0\beta^2\,|\omega_0|\,\Delta\omega(\Delta q)^3}\ln\left(1+\frac{|\omega_0|}{q_0}\right)\gg 1 \tag{IV.7.26}$$

the decoherence functional corresponding to two values of $\widetilde{\langle\ell\varphi\rangle}$ separated by $\widetilde{\langle\ell\Delta\varphi\rangle}$ will be small. This limit corresponds to

$$\left|\widetilde{\langle\ell\Delta\varphi\rangle}\right|\gg\frac{\beta\sqrt{8q_0\,|\omega_0|\,\Delta\omega(\Delta q)^3}}{(q_0^2-\omega_0^2)\sqrt{3\pi}}\left[\ln\left(1+\frac{|\omega_0|}{q_0}\right)\right]^{-1/2}; \tag{IV.7.27}$$

Considering the static limit $|\omega_0|\ll q_0$ for simplicity, (IV.7.27) becomes

$$\left|\widetilde{\langle\ell\Delta\varphi\rangle}\right|\gg\frac{\beta\sqrt{8q_0\,|\omega_0|\,\Delta\omega(\Delta q)^3}}{q_0^2\sqrt{3\pi}}\left[\frac{|\omega_0|}{q_0}\right]^{-1/2}=\beta q_0\sqrt{\frac{8\Delta\omega(\Delta q)^3}{3\pi q_0^4}}. \tag{IV.7.28}$$

For sufficiently small $\Delta\omega$ and Δq (which corresponds to averaging over a large spacetime region), the right hand side of (IV.7.28) becomes small, and thus (IV.7.28) can hold even in the perturbative limit where $\left|\widetilde{\langle\ell\Delta\varphi\rangle}\right|$ is small. [Recall that $\ell\Delta\varphi(\mathbf{x},t)$ is the representation in this toy model of a metric perturbation γ_{ab}.]

So a coarse graining which should decohere is one consisting of a set of alternatives $\{c_n\}$ which correspond to $\widetilde{\langle\ell\varphi\rangle}\in[n\Delta,(n+1)\Delta)$. As long as the size Δ of the regions is much larger than

$$\delta=\beta q_0\sqrt{\frac{8\Delta\omega(\Delta q)^3}{3\pi q_0^4}}, \tag{IV.7.29}$$

the analysis of Fig. III.1, with f replaced by $\widetilde{\langle\ell\varphi\rangle}$ and $G(f-f')$ replaced by e^{iW}, carries through, and off-diagonal elements of the decoherence functional $D(n_1,n_2)$ will be suppressed.

IV.8 Conclusions

In this chapter I have demonstrated that, in a simplified theory analogous to perturbative GR, some coarse grainings which restrict only the long wavelength modes of the field should decohere. This was done by calculating the influence functional e^{iW} between pairs of LWM histories which describes the effect of tracing out the short-wavelength modes. Decoherence is expected when $\left|e^{iW[\Phi_1,\Phi_2]}\right|$ becomes small for sufficiently large $\Phi_1 - \Phi_2$.

Even though the zero order term in the influence functional is unity, and one might normally assume that perturbative corrections cannot make $\left|e^{iW}\right|$ much smaller than one, it was possible to consider a regime where perturbation theory broke down enough to give $\left|e^{iW}\right| \ll 1$, but not so much that we were unable to calculate anything. This was done by working in the high-temperature regime where the inverse temperature β of the thermal state describing the SWMs was small. In this case terms which were higher order in the coupling ℓ could still become large for high temperature if they were proportional to, for example, $(\ell/\beta)^2$.

These ℓ/β terms in the influence functional were handled non-perturbatively for the terms in the action which are quadratic or linear in the SWMs, but the cubic terms in the action were analyzed using a perturbative expansion. That expansion showed that while there are corrections which go like $\frac{\mathcal{O}(\ell)}{\mathcal{O}(\ell)+\mathcal{O}(\beta)}$ or $\frac{\mathcal{O}(\beta)}{\mathcal{O}(\ell)+\mathcal{O}(\beta)}$, those are at largest $\mathcal{O}(1)$, and there are no $\mathcal{O}(\ell/\beta)$ terms to cancel out the effect from the quadratic action.

The reliance on perturbative analysis is one of the limitations of this result. It means that we can only analyze the question of decoherence in the high-temperature limit, defined in terms of the momentum k_c which defines SWMs from LWMs by $\beta k_c \ll 1$. If the temperature of the SWM thermal state is taken to be that of the present-day cosmic graviton background, the length scale corresponding to this limit is on the order of a millimeter.

Another problem comes from the non-renormalizability of our derivative action (a property it shares with GR itself). While the terms in the influence functional proportional to $(\ell/\beta)^2$ are finite, there are terms proportional to ℓ^2 alone which are ultraviolet divergent. We were able to ignore those by working in the high-temperature limit, but they may provide another way in which perturbation theory breaks down, demanding a fully non-perturbative analysis.

Before moving to a possible non-perturbative analysis, perhaps using the Regge calculus to skeletonize geometry, another improvement of this work would be to restore the tensor nature of the gravitational field and see if that modifies our scalar-field result.

And finally, the focus of this model has not been on cosmological systems (as contrasted to the matter-induced decoherence of spacetime described in [26] and [27]). The background spacetime was taken as Minkowski space and the temperature of the short-wavelength graviton state was taken to be its present-day value. Different background spacetimes might also be studied once the tensor nature of perturbative GR is restored.

IV.A Appendix: Expanding the action for perturbative GR to third order

In this appendix, I consider a metric

$$g_{ab}(\lambda) = g_{ab} + \lambda\gamma_{ab} \tag{IV.3.1}$$

and use an approach similar to that in Sec. 7.5 of [14] to expand the gravitational action constructed from $g_{ab}(\lambda)$

$$S = \frac{1}{16\pi G}\int \sqrt{|g(\lambda)|}\, d^4x R(\lambda), \tag{IV.A.1}$$

viewed as a function of γ_{ab} on the background spacetime with metric g_{ab}, in powers of λ out to third order.

The "scalar curvature" $R(\lambda)$ is constructed from the corresponding "Riemann tensor" ${R_{abc}}^d$ according to

$$R(\lambda) = g^{ac}(\lambda){R_{abc}}^b(\lambda) \tag{IV.A.2}$$

and has an expansion

$$R(\lambda) = \sum_n \lambda^n\, {}^nR. \tag{IV.A.3}$$

The "inverse metric" $g^{ab}(\lambda)$ is defined by[8]

$$g^{ab}(\lambda)g_{bc}(\lambda) = \delta^a_c \tag{IV.A.4}$$

and its expansion is given iteratively by

$${}^0g^{ab} = g^{ab} \tag{IV.A.5a}$$

$${}^{n+1}g^{ab} = -\,{}^ng^{ac}\gamma_{cd}g^{db} = -\,{}^ng^{ac}\gamma_c^{\,b}. \tag{IV.A.5b}$$

To work towards an expression for the curvature tensor ${R_{abc}}^d(\lambda)$ constructed from $g_{ab}(\lambda)$, one first considers the covariant derivative $\nabla_a(\lambda)$ constructed from $g_{ab}(\lambda)$ so that $\nabla_a(\lambda)g_{bc}(\lambda) = 0$. This is related to the covariant derivative ∇_a for the background metric by

$$\nabla_a(\lambda)\omega_b = \nabla_a\omega_b - C^c_{ab}(\lambda)\omega_c. \tag{IV.A.6}$$

$C^c_{ab}(\lambda)$ is called the connection tensor relating the two metrics, and in a particular coördinate system, it relates the Christoffel symbols for the two metrics:

$$C^\sigma_{\mu\nu}(\lambda) = \Gamma^\sigma_{\mu\nu}(\lambda) - \Gamma^\sigma_{\mu\nu}. \tag{IV.A.7}$$

[8] Note that it is *not* defined by raising the indices with g^{ab}: $g^{ab}(\lambda) \neq g^{ac}g_{cd}(\lambda)g^{db}$. We *are* using the convention that g_{ab} is the "actual" metric used in manipulating geometric quantities (for example $\gamma^{ab} = g^{ac}\gamma_{cd}g^{db}$), but that function is "turned off" for parametrized quantities [like $g_{ab}(\lambda)$] and their expansion terms (like ${}^ng_{ab}$), and such tensors with indices in new places must be explicitly defined as in (IV.A.4) or (IV.A.3).

It is given by

$$C^c_{ab}(\lambda) = \frac{1}{2} g^{cd}(\lambda)[\nabla_a g_{db}(\lambda) + \nabla_b g_{ad}(\lambda) - \nabla_d g_{ab}(\lambda)] = \frac{\lambda}{2} g^{cd}(\lambda)[\nabla_a \gamma_{db} + \nabla_b \gamma_{ad} - \nabla_d \gamma_{ab}]. \tag{IV.A.8}$$

Its expansion is given by

$${}^n C^c_{ab}(\lambda) = \frac{{}^{n-1}g^{cd}}{2}[\nabla_a \gamma_{db} + \nabla_b \gamma_{ad} - \nabla_d \gamma_{ab}]. \tag{IV.A.9}$$

where ${}^n g^{ab}$ is as given by (IV.A.5). In particular,

$${}^1 C^c_{ab} = \frac{1}{2}(\nabla_a \gamma^c_b + \nabla_b \gamma^c_a - \nabla^c \gamma_{ab}) \tag{IV.A.10a}$$

$${}^2 C^c_{ab} = -\frac{\gamma^{cd}}{2}(\nabla_a \gamma_{db} + \nabla_b \gamma_{ad} - \nabla_d \gamma_{ab}) \tag{IV.A.10b}$$

By using[9] $2\nabla_{[a}(\lambda)\nabla_{b]}(\lambda)\omega_c = R_{abc}{}^d(\lambda)\omega_d$ [see (I.B.7)] one has

$$R_{abc}{}^d(\lambda) = R_{abc}{}^d - 2\nabla_{[a} C^d_{b]c}(\lambda) + 2C^e_{c[a}(\lambda) C^d_{b]e}(\lambda), \tag{IV.A.11}$$

which contracts, recalling that g_{ab} is Ricci-flat, to

$$R_{ab}(\lambda) = -2\nabla_{[a} C^c_{c]b}(\lambda) + 2C^d_{b[a}(\lambda) C^c_{c]d}(\lambda). \tag{IV.A.12}$$

This then gives the expansion

$${}^0 R_{ab} = 0 \tag{IV.A.13a}$$

$${}^1 R_{ab} = -2\nabla_{[a}\, {}^1 C^c_{c]b} \tag{IV.A.13b}$$

$${}^2 R_{ab} = -2\nabla_{[a}\, {}^2 C^c_{c]b} + 2\, {}^1 C^d_{b[a}\, {}^1 C^c_{c]d} \tag{IV.A.13c}$$

$${}^3 R_{ab} = -2\nabla_{[a}\, {}^3 C^c_{c]b} + 2\, {}^1 C^d_{b[a}\, {}^2 C^c_{c]d} + 2\, {}^2 C^d_{b[a}\, {}^1 C^c_{c]d}. \tag{IV.A.13d}$$

The scalar curvature is given by inserting (IV.A.13) and into [*cf.* (IV.A.5)]

$${}^0 R = 0 \tag{IV.A.14a}$$

$${}^1 R = g^{ab}\, {}^1 R_{ab} \tag{IV.A.14b}$$

$${}^2 R = g^{ab}\, {}^2 R_{ab} - \gamma^{ab}\, {}^1 R_{ab} \tag{IV.A.14c}$$

$${}^3 R = g^{ab}\, {}^3 R_{ab} - \gamma^{ab}\, {}^2 R_{ab} + \gamma^{ac}\gamma_c^b\, {}^1 R_{ab}. \tag{IV.A.14d}$$

Of course, it is not enough simply to expand $R(\lambda)$; the action (IV.A.1) also includes a factor of $\sqrt{|g(\lambda)|}$. Now, since

$$g_{ab}(\lambda) = g_{ac}(\delta^c_b + \lambda \gamma^c_b), \tag{IV.A.15}$$

[9] I introduce here the usual antisymmetrization notation $\omega_{[ab]} = \frac{1}{2}(\omega_{ab} - \omega_{ba})$ *etc.*

the determinant is given by

$$g(\lambda) = g[1 + \lambda\gamma + \mathcal{O}(\lambda^2)], \tag{IV.A.16}$$

where $\gamma = g^{ab}\gamma_{ab}$ is the trace of γ_{ab}. This means we can easily expand $\sqrt{|g(\lambda)|}$ out to first order:

$$^0\sqrt{|g|} = \sqrt{|g|} \tag{IV.A.17a}$$

$$^1\sqrt{|g|} = \frac{\gamma}{2}\sqrt{|g|} \tag{IV.A.17b}$$

to get the expansion for the action

$$^0S = 0 \tag{IV.A.18a}$$

$$^1S = \frac{1}{16\pi G}\int \sqrt{|g|}\, d^4x\, {}^1R \tag{IV.A.18b}$$

$$^2S = \frac{1}{16\pi G}\int \sqrt{|g|}\, d^4x \left({}^2R + \frac{\gamma}{2}\, {}^1R \right) \tag{IV.A.18c}$$

$$^3S = \frac{1}{16\pi G}\int \sqrt{|g|}\, d^4x \left({}^3R + \frac{\gamma}{2}\, {}^2R + \frac{{}^2(\sqrt{|g|})}{\sqrt{|g|}}\, {}^1R \right). \tag{IV.A.18d}$$

There would appear to be some difficulty in expanding to third order, as (IV.A.18d) contains the as yet unknown ${}^2(\sqrt{|g|})$. This is resolved by the detailed form of ${}^1R$; to see this, though, we must consider the issue of the diffeomorphism invariance of S in the next section.

IV.A.1 Gauge freedom

If we make the infinitesimal transformation [*cf.* (I.B.24)]

$$g_{ab}(\lambda) \to \widetilde{g}_{ab}(\lambda) = g_{ab}(\lambda) - 2\lambda\nabla_{(a}\delta x_{b)}, \tag{IV.A.19}$$

the geometry described by $\widetilde{g}_{ab}(\lambda)$ is the same as that described by $g_{ab}(\lambda)$. Since $\widetilde{g}_{ab}(0) = g_{ab} = g_{ab}(0)$, we can define

$$\widetilde{\gamma}_{ab} = \widetilde{g}_{ab}(1) - g_{ab} \tag{IV.A.20}$$

so that

$$\widetilde{g}_{ab}(\lambda) = g_{ab} + \lambda\widetilde{\gamma}_{ab} \tag{IV.A.21}$$

This means that the perturbations γ_{ab} and $\widetilde{\gamma}_{ab}$ on the same background metric g_{ab} give the same perturbed geometry, so γ_{ab} has a gauge symmetry

$$\gamma_{ab} \to \widetilde{\gamma}_{ab} = \gamma_{ab} - 2\lambda\nabla_{(a}\delta x_{b)}. \tag{IV.A.22}$$

Building up finite transforms from the infinitesimal (IV.A.22), one can [14] convert any perturbation γ_{ab}, by a suitable gauge transformation, into the transverse, traceless (TT) gauge in which

$$\nabla_a\gamma^a_b = 0 \tag{IV.A.23a}$$

$$\gamma = 0. \tag{IV.A.23b}$$

IV.A.2 Expansion terms in the transverse, traceless gauge

In the TT gauge, we have from (IV.A.10a)

$$ {}^1C^c_{cb} = \frac{1}{2}\nabla_b\gamma = 0 \tag{IV.A.24}$$

so

$$ {}^1R_{ab} = \nabla_c\, {}^1C^c_{ab} = \nabla_c\nabla_{(a}\gamma^c_{b)} - \frac{1}{2}\nabla_c\nabla^c\gamma_{ab} \tag{IV.A.25}$$

Contracting a and b makes the first term proportional to $\nabla_a\gamma^{ca}$ and the second proportional to γ, so ${}^1R$ vanishes in the TT gauge.[10] This means that (IV.A.18) reduces in this gauge to

$$ {}^nS = \begin{cases} 0 & n = 0 \text{ or } 1 \\ \frac{1}{16\pi G}\int\sqrt{|g|}\,d^4x\,{}^nR & n = 2 \text{ or } 3 \end{cases} \tag{IV.A.26}$$

Using (I.B.7), we have the result that in the TT gauge

$$ \nabla_c\nabla_a\gamma^c_b = \nabla_a\nabla_c\gamma^c_b + R_{cab}{}^d\gamma^c_d - R_{cad}{}^c\gamma^d_b = R_{cab}{}^d\gamma^c_d, \tag{IV.A.27}$$

where we have used the fact that $R_{ab} = 0$. Thus

$$ {}^1R_{ab} = -\frac{1}{2}\nabla_c\nabla^c\gamma_{ab} - R_{acbd}\gamma^{cd}, \tag{IV.A.28}$$

Note that this first order expansion gives the first order vacuum Einstein equation (7.5.23) in [14].

The second order expansion term ${}^2R_{ab}$ is given by (IV.A.13c). In the TT gauge, this becomes

$$ {}^2R_{ab} = -2\nabla_{[a}\,{}^2C^c_{c]b} - {}^1C^d_{bc}\,{}^1C^c_{ad}. \tag{IV.A.29}$$

For the calculation of ${}^2S$, the divergence term will contribute only a surface term, so we put it aside for the moment and consider

$$\begin{aligned} {}^2R_{ab} + 2\nabla_{[a}\,{}^2C^c_{c]b} &= -\left(\nabla_{(a}\gamma^c_{d)} - \frac{1}{2}\nabla^c\gamma_{ad}\right)\left(\nabla_{(b}\gamma^d_{c)} - \frac{1}{2}\nabla^d\gamma_{bc}\right) \\ &= -\frac{1}{4}(\nabla_a\gamma^c_d)(\nabla_b\gamma^d_c) + \frac{1}{2}(\nabla_c\gamma^d_a)(\nabla^c\gamma_{bd}) - \frac{1}{2}(\nabla_c\gamma^d_a)(\nabla_d\gamma^c_b) \\ &= -\frac{1}{4}(\nabla_a\gamma^c_d)(\nabla_b\gamma^d_c) + \frac{1}{2}(\nabla_c\gamma^d_a)(\nabla^c\gamma_{bd}) - \frac{1}{2}\nabla_c(\gamma^d_a\nabla_d\gamma^c_b) + \frac{1}{2}\gamma^d_a R_{cdb}{}^e\gamma^c_e \end{aligned} \tag{IV.A.30}$$

Thus the second order action is, ignoring surface terms,

$$\begin{aligned} {}^2S =& \frac{1}{16\pi G}\int\sqrt{|g|}\,d^4x(g^{ab}\,{}^2R_{ab} - \gamma^{ab}\,{}^1R_{ab}) \\ =& \frac{1}{16\pi G}\int\sqrt{|g|}\,d^4x\left[-\frac{1}{4}(\nabla_c\gamma_{ab})(\nabla^c\gamma^{ab}) + \frac{1}{2}\gamma^{ab}R_{acbd}\gamma^{cd}\right]. \end{aligned} \tag{IV.A.31}$$

[10] Note that in *any* gauge, ${}^1R = \nabla_a(-g^{ab}\,{}^1C^c_{cb} + g^{cb}\,{}^1C^a_{cb})$ is a total divergence, so ${}^1S$ is just a surface term. However, the TT gauge result that ${}^1R = 0$ allows us to remove it from ${}^2S$ and ${}^3S$ as well.

Varying this with respect to γ^{ab} gives the first order equation ${}^1R_{ab} = 0$.

Finally, we calculate the third order action ${}^3S = (16\pi G)^{-1} \int \sqrt{|g|}\, d^4x\, {}^3R$. First, note that all the terms in ${}^3R$ can be brought, by integration by parts and relabelling of indices, into one of the following forms:

$$\gamma^{ab}(\nabla_c \gamma_a^d)(\nabla^c \gamma_{bd}), \quad \gamma^{ab}(\nabla_a \gamma_c^d)(\nabla_b \gamma_d^c), \quad \gamma^{ab}(\nabla_c \gamma_a^d)(\nabla_d \gamma_b^c), \quad \text{or} \quad \gamma^{ab} R_{acb}{}^d \gamma^{ce} \gamma_{de}. \tag{IV.A.32}$$

There is one other term which is allowed by dimensional analysis and the TT gauge conditions, but it is related to these by

$$\begin{aligned} \gamma^{ab}(\nabla_a \gamma_c^d)(\nabla_d \gamma_b^c) \cong & -(\nabla_d \gamma^{ab})(\nabla_a \gamma_c^d)\gamma_b^c - \gamma^{ab}(\nabla_d \nabla_a \gamma_c^d)\gamma_b^c \\ = & -\gamma^{ab}(\nabla_c \gamma_a^d)(\nabla_d \gamma_b^c) + \gamma^{ab} R_{acb}{}^d \gamma^{ce}\gamma_{de}, \end{aligned} \tag{IV.A.33}$$

Where $\cong$ indicates equality up to a total divergence. Now,

$$g^{ab}\,{}^3R_{ab} = g^{ab}(-2\nabla_{[a}\,{}^3C^c_{c]b} + 2\,{}^1C^d_{b[a}\,{}^2C^c_{c]d} + 2\,{}^2C^d_{b[a}\,{}^1C^c_{c]d}) \tag{IV.A.34}$$

is simplified by noting that the first term is a total derivative, that ${}^1C^c_{ab} g^{ab} = 0$ in the TT gauge, and that ${}^2C^c_{ab} = -\gamma^c_d\,{}^1C^d_{ab}$, so

$$\begin{aligned} g^{ab}\,{}^3R_{ab} \cong & 2g^{ab}\,{}^1C^d_{bc}\,{}^1C^e_{ad}\gamma^c_e = \frac{\gamma^{ab}}{2}(\nabla_a\gamma_c^d + \nabla_c\gamma_a^d - \nabla^d\gamma_{ac})(\nabla^c\gamma_{bd} + \nabla_d\gamma_b^c - \nabla_b\gamma_d^c) \\ = & \frac{\gamma^{ab}}{2}[-(\nabla_a\gamma_c^d)(\nabla_b\gamma_d^c) + 2(\nabla_a\gamma_c^d)(\nabla_d\gamma_b^c)] \\ \cong & \gamma^{ab}\left[-\frac{1}{2}(\nabla_a\gamma_c^d)(\nabla_b\gamma_d^c) - (\nabla_c\gamma_a^d)(\nabla_d\gamma_b^c) + R_{acb}{}^d\gamma^{ce}\gamma_{de}\right]. \end{aligned} \tag{IV.A.35}$$

Now we move on to $-\gamma^{ab}\,{}^2R_{ab}$, using (IV.A.30) to say

$$\begin{aligned} {}^2R_{ab} = -\nabla_a\,{}^2C^c_{cb} + \nabla_c\left({}^2C^c_{ab} - \frac{1}{2}\gamma_a^d\nabla_d\gamma_b^c\right) \\ - \frac{1}{4}(\nabla_a\gamma_d^c)(\nabla_b\gamma_c^d) + \frac{1}{2}(\nabla_c\gamma_a^d)(\nabla^c\gamma_{bd}) + \frac{1}{2}\gamma_a^d R_{cdb}{}^e\gamma_e^c; \end{aligned} \tag{IV.A.36}$$

since

$${}^2C^c_{ab} - \frac{1}{2}\gamma_a^d\nabla_d\gamma_b^c = -\gamma^{cd}\nabla_{(a}\gamma_{b)d} + \frac{1}{2}\gamma^{cd}\nabla_d\gamma_{ab} - \frac{1}{2}\gamma_a^d\nabla_d\gamma_b^c, \tag{IV.A.37}$$

we have

$$
\begin{aligned}
-\gamma^{ab}\,{}^2R_{ab} \cong& (\nabla_c\gamma^{ab})\left(-\gamma^{cd}\nabla_a\gamma_{bd} + \frac{1}{2}\gamma^{cd}\nabla_d\gamma_{ab} - \frac{1}{2}\gamma_a^d\nabla_d\gamma_b^c\right)\\
&- \gamma^{ab}\left[-\frac{1}{4}(\nabla_a\gamma_d^c)(\nabla_b\gamma_c^d) + \frac{1}{2}(\nabla_c\gamma_a^d)(\nabla^c\gamma_{bd}) - \frac{1}{2}R_{acb}{}^d\gamma^{ce}\gamma_{de}\right]\\
=&\gamma^{ab}\left[-\frac{3}{2}(\nabla_a\gamma_c^d)(\nabla_d\gamma_b^c) + \frac{3}{4}(\nabla_a\gamma_d^c)(\nabla_b\gamma_c^d) - \frac{1}{2}(\nabla_c\gamma_a^d)(\nabla^c\gamma_{bd}) + \frac{1}{2}R_{acb}{}^d\gamma^{ce}\gamma_{de}\right]\\
\cong&\gamma^{ab}\left[-\frac{1}{2}(\nabla_c\gamma_a^d)(\nabla^c\gamma_{bd}) + \frac{3}{4}(\nabla_a\gamma_d^c)(\nabla_b\gamma_c^d) + \frac{3}{2}(\nabla_c\gamma_a^d)(\nabla_d\gamma_b^c) - R_{acb}{}^d\gamma^{ce}\gamma_{de}\right]
\end{aligned}
\tag{IV.A.38}
$$

Finally,

$$
\gamma^{ac}\gamma_c^b\,{}^1R_{ab} = -\frac{1}{2}\gamma^{ac}\gamma_c^b\nabla_d\nabla^d\gamma_{ab} - \gamma^{ac}\gamma_c^bR_{aeb}{}^d\gamma_d^e \cong \gamma^{ab}\left[(\nabla_c\gamma_a^d)(\nabla^c\gamma_{bd}) - R_{acb}{}^d\gamma^{ce}\gamma_{de}\right], \tag{IV.A.39}
$$

so

$$
{}^3R \cong \gamma^{ab}\left[\frac{1}{2}(\nabla_c\gamma_a^d)(\nabla^c\gamma_{bd}) + \frac{1}{4}(\nabla_a\gamma_c^d)(\nabla_b\gamma_d^c) + \frac{1}{2}(\nabla_c\gamma_a^d)(\nabla_d\gamma_b^c) - R_{acb}{}^d\gamma^{ce}\gamma_{de}\right] \tag{IV.A.40}
$$

So the action to third order in λ, not including surface terms, is

$$
\begin{aligned}
S = \frac{1}{16\pi G}\int\sqrt{|g|}\,d^4x\Big\{&\lambda^2\left[-\frac{1}{4}(\nabla_c\gamma_{ab})(\nabla^c\gamma^{ab}) + \frac{1}{2}\gamma^{ab}R_{acbd}\gamma^{cd}\right]\\
&\lambda^3\gamma^{ab}\left[\frac{1}{2}(\nabla_c\gamma_a^d)(\nabla^c\gamma_{bd}) + \frac{1}{4}(\nabla_a\gamma_c^d)(\nabla_b\gamma_d^c) + \frac{1}{2}(\nabla_c\gamma_a^d)(\nabla_d\gamma_b^c) - R_{acb}{}^d\gamma^{ce}\gamma_{de}\right]\\
&+\mathcal{O}(\lambda^4)\Big\}.
\end{aligned}
\tag{IV.A.41}
$$

IV.B Appendix: The propagator for a time-dependent harmonic oscillator

For a simple harmonic oscillator with one degree of freedom $\phi(t)$ and the lagrangian

$$
L(t) = \frac{1}{2}\left[m\dot{\phi}^2(t) - m\omega^2\phi^2(t)\right] \tag{IV.B.1}
$$

one can easily calculate the propagator via a path integral[41]:

$$
\begin{aligned}
K(\phi_bt_b|\phi_at_a) =& \int\limits_{\phi_b\phi_a}\mathcal{D}\phi\exp\left[\frac{i}{2}\int_{t_a}^{t_b}dt(m\dot{\phi}^2 - m\omega^2\phi^2)\right]\\
=&\sqrt{\frac{m\omega}{2\pi i\sin\omega T_{ba}}}\exp\left[\frac{im\omega}{2\sin\omega T_{ba}}\begin{pmatrix}\phi_b\\ \phi_a\end{pmatrix}^{\mathrm{tr}}\begin{pmatrix}\cos\omega T_{ba} & -1\\ -1 & \cos\omega T_{ba}\end{pmatrix}\begin{pmatrix}\phi_b\\ \phi_a\end{pmatrix}\right].
\end{aligned}
\tag{IV.B.2}
$$

The purpose of this appendix is to calculate the propagator when the coördinate ϕ is a real vector, and the lagrangian is

$$L(t) = \frac{1}{2}\left[\dot{\phi}(t)^{\mathrm{tr}} m(t)\dot{\phi}(t) - \phi(t)^{\mathrm{tr}}\varpi(t)\phi(t)\right], \tag{IV.B.3}$$

where the arbitrary time-dependent real symmetric matrices $m(t)$ and $\varpi(t)$ have taken the place of the constants m and $m\omega^2$, respectively.

First, for the case of time-independent matrices m and ϖ, we observe that as the action is quadratic, the propagator must have the form[41] $K = F(T_{ba})e^{iS_{\mathrm{cl}}}$ (where $T_{ba} = t_b - t_a$). The action S_{cl} of the classical path with endpoints $\phi_{\mathrm{cl}}(t_a) = \phi_a$ and $\phi_{\mathrm{cl}}(t_b) = \phi_b$ can be calculated by defining $q = m^{1/2}\phi$ so that

$$L(t) = \frac{1}{2}\left[\dot{q}(t)^{\mathrm{tr}}\dot{q}(t) - q(t)^{\mathrm{tr}} m^{-1/2}\varpi m^{-1/2} q(t)\right]. \tag{IV.B.4}$$

Defining Ω by $\Omega^2 = m^{-1/2}\varpi m^{-1/2}$, the classical solution is

$$q_{\mathrm{cl}}(t) = \cos\Omega(t-t_a)q_a + \sin\Omega(t-t_a)\left(\frac{1}{\sin\Omega T_{ba}}q_b - \frac{\cos\Omega T_{ba}}{\sin\Omega T_{ba}}q_a\right), \tag{IV.B.5}$$

which makes the classical action

$$\begin{aligned} S_{\mathrm{cl}} &= \frac{1}{2}\int_{t_a}^{t_b} dt(\dot{q}_{\mathrm{cl}}^{\mathrm{tr}}\dot{q}_{\mathrm{cl}} - q_{\mathrm{cl}}{}^{\mathrm{tr}}\Omega^2 q_{\mathrm{cl}}) = \frac{1}{2}\begin{pmatrix} q_b \\ q_a \end{pmatrix}^{\mathrm{tr}} \begin{pmatrix} \frac{\Omega\cos\Omega T_{ba}}{\sin\Omega T_{ba}} & -\frac{\Omega}{\sin\Omega T_{ba}} \\ -\frac{\Omega}{\sin\Omega T_{ba}} & \frac{\Omega\cos\Omega T_{ba}}{\sin\Omega T_{ba}} \end{pmatrix} \begin{pmatrix} q_b \\ q_a \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} \phi_b \\ \phi_a \end{pmatrix}^{\mathrm{tr}} m^{1/2} \begin{pmatrix} \frac{\Omega\cos\Omega T_{ba}}{\sin\Omega T_{ba}} & -\frac{\Omega}{\sin\Omega T_{ba}} \\ -\frac{\Omega}{\sin\Omega T_{ba}} & \frac{\Omega\cos\Omega T_{ba}}{\sin\Omega T_{ba}} \end{pmatrix} m^{1/2} \begin{pmatrix} \phi_b \\ \phi_a \end{pmatrix}. \end{aligned} \tag{IV.B.6}$$

This tells us that the infinitesimal propagator in the time-dependent case is

$$\begin{aligned} &K(\phi_b t + dt|\phi_a t) \\ &\quad = F(t+dt|t)\exp\left[\frac{i}{2}\begin{pmatrix} \phi_b \\ \phi_a \end{pmatrix}^{\mathrm{tr}} m^{1/2}(t) \begin{pmatrix} \frac{1-\Omega^2(t)(dt)^2/2}{dt} & -\frac{1}{dt} \\ -\frac{1}{dt} & \frac{1-\Omega^2(t)(dt)^2/2}{dt} \end{pmatrix} m^{1/2}(t) \begin{pmatrix} \phi_b \\ \phi_a \end{pmatrix}\right] \\ &\quad = F(t+dt|t)\exp\left[\frac{i}{2}\begin{pmatrix} \phi_b \\ \phi_a \end{pmatrix}^{\mathrm{tr}} \begin{pmatrix} \frac{m(t)-\varpi(t)(dt)^2/2}{dt} & -\frac{m(t)}{dt} \\ -\frac{m(t)}{dt} & \frac{m(t)-\varpi(t)(dt)^2/2}{dt} \end{pmatrix} \begin{pmatrix} \phi_b \\ \phi_a \end{pmatrix}\right] \end{aligned} \tag{IV.B.7}$$

We can then build up the finite propagator $K_{ba} = K(\phi_b t_b|\phi_a t_a)$ by the composition rule $K_{ba} = \int d\phi_c K_{bc}K_{ca}$ [which also allows us to determine the coëfficient$F(t_b|t_a)$]. A little tinkering reveals that the form for K which is preserved under composition is

$$K_{ba} = \frac{1}{\sqrt{\det(2\pi i\mathfrak{C}_{ba})}}\exp\left[\frac{i}{2}\begin{pmatrix} \phi_b \\ \phi_a \end{pmatrix}^{\mathrm{tr}} \begin{pmatrix} \mathfrak{C}_{ba}^{-1}\mathfrak{B}_{ba} & -\mathfrak{C}_{ba}^{-1} \\ -\mathfrak{C}_{ba}^{-1\,\mathrm{tr}} & \mathfrak{A}_{ba}\mathfrak{C}_{ba}^{-1} \end{pmatrix} \begin{pmatrix} \phi_b \\ \phi_a \end{pmatrix}\right], \tag{IV.B.8}$$

where $\mathfrak{A}_{ba} = \mathfrak{A}(t_b|t_a)$, $\mathfrak{B}_{ba} = \mathfrak{B}(t_b|t_a)$, and $\mathfrak{C}_{ba} = \mathfrak{C}(t_b|t_a)$ are all real matrices to be determined. In the infinitesimal case, they are given from (IV.B.7) by

$$\mathfrak{A}(t+dt|t) = 1 - \varpi(t)m^{-1}(t)\frac{dt^2}{2} + \mathcal{O}(dt^4) \tag{IV.B.9a}$$

$$\mathfrak{B}(t+dt|t) = 1 - m^{-1}(t)\varpi(t)\frac{dt^2}{2} + \mathcal{O}(dt^4) \tag{IV.B.9b}$$

$$\mathfrak{C}(t+dt|t) = m^{-1}(t)dt + \mathcal{O}(dt^3) \tag{IV.B.9c}$$

Without loss of generality, we can take the matrix in the exponential to be symmetric, which makes $\mathfrak{A}\mathfrak{C}^{-1} = {\mathfrak{C}^{-1}}^{\mathrm{tr}}\mathfrak{A}^{\mathrm{tr}}$ and $\mathfrak{C}^{-1}\mathfrak{B} = \mathfrak{B}^{\mathrm{tr}}{\mathfrak{C}^{-1}}^{\mathrm{tr}}$.[11] Using the Gaussian integral

$$\int dv\, e^{\frac{i}{2}v^{\mathrm{tr}}Av \pm ib^{\mathrm{tr}}v} = \frac{e^{-\frac{i}{2}b^{\mathrm{tr}}A^{-1}b}}{\sqrt{\det\left(\frac{A}{2\pi i}\right)}}, \tag{IV.B.10}$$

we can perform the composition

$$\begin{aligned}
K_{ba} &= \int d\phi_c K_{bc}K_{ca} \\
&= \frac{1}{\sqrt{\det(2\pi i\mathfrak{C}_{bc})\det(2\pi i\mathfrak{C}_{ca})}}\exp\left\{\frac{i}{2}\left[{\phi_c}^{\mathrm{tr}}(\mathfrak{A}_{bc}\mathfrak{C}_{bc}^{-1} + \mathfrak{C}_{ca}^{-1}\mathfrak{B}_{ca})\phi_c\right.\right. \\
&\qquad \left.\left. -2{\phi_c}^{\mathrm{tr}}({\mathfrak{C}_{bc}^{-1}}^{\mathrm{tr}}\phi_b + \mathfrak{C}_{ca}^{-1}\phi_a) + {\phi_b}^{\mathrm{tr}}\mathfrak{C}_{bc}^{-1}\mathfrak{B}_{bc}\phi_b + {\phi_a}^{\mathrm{tr}}\mathfrak{A}_{ca}\mathfrak{C}_{ca}^{-1}\phi_a\right]\right\} \\
&= \frac{1}{\sqrt{\det(2\pi i\mathfrak{C}_{bc})\det\left(\frac{\mathfrak{A}_{bc}\mathfrak{C}_{bc}^{-1}+\mathfrak{C}_{ca}^{-1}\mathfrak{B}_{ca}}{2\pi i}\right)\det(2\pi i\mathfrak{C}_{ca})}}\exp\left\{\frac{i}{2}\left[{\phi_b}^{\mathrm{tr}}\mathfrak{C}_{bc}^{-1}\mathfrak{B}_{bc}\phi_b\right.\right. \\
&\quad \left.\left. +{\phi_a}^{\mathrm{tr}}\mathfrak{A}_{ca}\mathfrak{C}_{ca}^{-1}\phi_a - ({\phi_b}^{\mathrm{tr}}\mathfrak{C}_{bc}^{-1} + {\phi_a}^{\mathrm{tr}}{\mathfrak{C}_{ca}^{-1}}^{\mathrm{tr}})(\mathfrak{A}_{bc}\mathfrak{C}_{bc}^{-1} + \mathfrak{C}_{ca}^{-1}\mathfrak{B}_{ca})^{-1}({\mathfrak{C}_{bc}^{-1}}^{\mathrm{tr}}\phi_b + \mathfrak{C}_{ca}^{-1}\phi_a)\right]\right\}
\end{aligned} \tag{IV.B.11}$$

and read off

$$\mathfrak{A}_{ba} = \mathfrak{A}_{ca}\mathfrak{A}_{bc} + \mathfrak{A}_{ca}\mathfrak{C}_{ca}^{-1}\mathfrak{B}_{ca}\mathfrak{C}_{bc} - {\mathfrak{C}_{ca}^{-1}}^{\mathrm{tr}}\mathfrak{C}_{bc} \tag{IV.B.12a}$$

$$\mathfrak{B}_{ba} = \mathfrak{B}_{ca}\mathfrak{B}_{bc} + \mathfrak{C}_{ca}\mathfrak{A}_{bc}\mathfrak{C}_{bc}^{-1}\mathfrak{B}_{bc} - \mathfrak{C}_{ca}{\mathfrak{C}_{bc}^{-1}}^{\mathrm{tr}} \tag{IV.B.12b}$$

$$\mathfrak{C}_{ba} = \mathfrak{C}_{ca}\mathfrak{A}_{bc} + \mathfrak{B}_{ca}\mathfrak{C}_{bc}. \tag{IV.B.12c}$$

We can use these, along with the values (IV.B.9) for infinitesimal propagation, to get expressions for $\mathfrak{A}(t_b+dt_b|t_a)$, $\mathfrak{B}(t_b+dt_b|t_a)$, and $\mathfrak{C}(t_b+dt_b|t_a)$ in terms of $\mathfrak{A}(t_b|t_a)$, *etc.*, which

[11] The matrix in (IV.B.8) seems to treat the matrices and their transposes unequally, but as $\begin{pmatrix} \mathfrak{C}^{-1}\mathfrak{B} & -\mathfrak{C}^{-1} \\ -{\mathfrak{C}^{-1}}^{\mathrm{tr}} & \mathfrak{A}\mathfrak{C}^{-1} \end{pmatrix} = \begin{pmatrix} \mathfrak{B}^{\mathrm{tr}}{\mathfrak{C}^{-1}}^{\mathrm{tr}} & -{\mathfrak{C}^{-1}}^{\mathrm{tr}} \\ -\mathfrak{C}^{-1} & {\mathfrak{C}^{-1}}^{\mathrm{tr}}\mathfrak{A}^{\mathrm{tr}} \end{pmatrix}$, this is not the case.

give the differential equations

$$\frac{d\mathfrak{A}(t_b|t_a)}{dt_b} = \left[\mathfrak{A}(t_b|t_a)\mathfrak{C}^{-1}(t_b|t_a)\mathfrak{B}(t_b|t_a) - \mathfrak{C}^{-1}(t_b|t_a)^{\mathrm{tr}}\right] m^{-1}(t_b) \tag{IV.B.13a}$$

$$\frac{d\mathfrak{B}(t_b|t_a)}{dt_b} = -\mathfrak{C}(t_b|t_a)\varpi(t_b) \tag{IV.B.13b}$$

$$\frac{d\mathfrak{C}(t_b|t_a)}{dt_b} = \mathfrak{B}(t_b|t_a)m^{-1}(t_b); \tag{IV.B.13c}$$

similarly, the expressions for $\mathfrak{A}(t_b|t_a + dt_a)$ give the equations

$$\frac{d\mathfrak{A}(t_b|t_a)}{dt_a} = \varpi(t_a)\mathfrak{C}(t_b|t_a) \tag{IV.B.14a}$$

$$\frac{d\mathfrak{B}(t_b|t_a)}{dt_a} = -m^{-1}(t_a)\left[\mathfrak{A}(t_b|t_a)\mathfrak{C}^{-1}(t_b|t_a)\mathfrak{B}(t_b|t_a) - \mathfrak{C}^{-1}(t_b|t_a)^{\mathrm{tr}}\right] \tag{IV.B.14b}$$

$$\frac{d\mathfrak{C}(t_b|t_a)}{dt_a} = -m^{-1}(t_a)\mathfrak{A}(t_b|t_a). \tag{IV.B.14c}$$

The initial conditions for either of these systems of equations are [see (IV.B.9)] $\mathfrak{A}(t|t) = 1 = \mathfrak{B}(t|t)$ and $\mathfrak{C}(t|t) = 0$.

Focusing on the t_b equations, we can write (IV.B.13b) and (IV.B.13c) as a larger matrix equation

$$\frac{d}{dt_b}\begin{pmatrix}\mathfrak{B}(t_b|t_a)^{\mathrm{tr}}\\ \mathfrak{C}(t_b|t_a)^{\mathrm{tr}}\end{pmatrix} = \begin{pmatrix}0 & -\varpi(t_b)\\ m^{-1}(t_b) & 0\end{pmatrix}\begin{pmatrix}\mathfrak{B}(t_b|t_a)^{\mathrm{tr}}\\ \mathfrak{C}(t_b|t_a)^{\mathrm{tr}}\end{pmatrix} = M(t_b)\begin{pmatrix}\mathfrak{B}(t_b|t_a)^{\mathrm{tr}}\\ \mathfrak{C}(t_b|t_a)^{\mathrm{tr}}\end{pmatrix} \tag{IV.B.15}$$

whose solution can be written in terms of a time-ordered exponential

$$\begin{pmatrix}\mathfrak{B}(t_b|t_a)^{\mathrm{tr}}\\ \mathfrak{C}(t_b|t_a)^{\mathrm{tr}}\end{pmatrix} = \mathrm{T}\exp\int_{t_a}^{t_b} dt M(t)\begin{pmatrix}1\\0\end{pmatrix} = \sum_{n=0}^{\infty}\left(\prod_{k=1}^{n}\int_{t_a}^{\tilde{t}_{k-1}} d\tilde{t}_k M(\tilde{t}_k)\right)\begin{pmatrix}1\\0\end{pmatrix} \tag{IV.B.16}$$

where $\tilde{t}_0 = t_b$. Since

$$M(t)M(\tilde{t}) = \begin{pmatrix}-\varpi(t)m^{-1}(\tilde{t}) & 0\\ 0 & -m^{-1}(t)\varpi(\tilde{t})\end{pmatrix}, \tag{IV.B.17}$$

$\mathfrak{B}^{\mathrm{tr}}$ contains only terms with even numbers of Ms, and

$$\mathfrak{B}(t_b|t_a)^{\mathrm{tr}} = \sum_{n=0}^{\infty}\prod_{k=1}^{n}\left(-\int_{t_a}^{\tilde{t}_{k-1}} dt_k \int_{t_a}^{t_k} d\tilde{t}_k \varpi(t_k)m^{-1}(\tilde{t}_k)\right). \tag{IV.B.18}$$

Thus

$$\mathfrak{B}(t_b|t_a) = \sum_{n=0}^{\infty}\left(\prod_{k=1}^{n}\int_{t_a}^{\tilde{t}_{k-1}} dt_k \int_{t_a}^{t_k} d\tilde{t}_k\right)\prod_{k=n}^{1}\left[-m^{-1}(\tilde{t}_k)\varpi(t_k)\right], \tag{IV.B.19a}$$

and we can obtain the other two matrices by using (IV.B.13b) and (IV.B.14c):

$$\mathfrak{C}(t_b|t_a) = -\frac{d\mathfrak{B}(t_b|t_a)}{dt_b}\varpi^{-1}(t_b) \tag{IV.B.19b}$$

$$\mathfrak{A}(t_b|t_a) = -m(t_a)\frac{d\mathfrak{C}(t_b|t_a)}{dt_a}. \tag{IV.B.19c}$$

IV.C Appendix: Perturbative expansions

IV.C.1 $\mathfrak{A}$, $\mathfrak{B}$ and $\mathfrak{C}$

To make practical use of the exact expressions (IV.B.19), we need to expand them in powers of the coupling constant λ. The zero-order expansion is trivial, since $m_0 = 1$ and $\varpi_0 = \Omega_0^2$, where

$$\Omega_0 = \begin{pmatrix} \{\delta_{\mathbf{M}_1\mathbf{M}_2} k_{\mathbf{M}_1}\} & \{0\} \\ \{0\} & \{\delta_{\mathbf{M}_1\mathbf{M}_2} k_{\mathbf{M}_1}\} \end{pmatrix}. \tag{IV.C.1}$$

The expansion is

$$\mathfrak{A}_0 = \cos\Omega_0 T_{ba} = \mathfrak{B}_0 \tag{IV.C.2a}$$

$$\mathfrak{C}_0 = \frac{\sin\Omega_0 T_{ba}}{\Omega_0} \tag{IV.C.2b}$$

Proceeding to the first order terms, we first need to expand

$$\begin{aligned}\prod_{k=n}^{1}\left[-m^{-1}(\tilde{t}_k)\varpi(t_k)\right] &= \prod_{k=n}^{1}\left\{-[1+\lambda m_1(\tilde{t}_k)]^{-1}[\Omega_0^2+\lambda\varpi_1(t_k)]\right\} \\ &= \prod_{k=n}^{1}[1+\lambda m_1(\tilde{t}_k)]^{-1}[1+\lambda\mu_1(t_k)](-\Omega_0^2),\end{aligned} \tag{IV.C.3}$$

where I have defined $\mu_1(t_k) = \varpi_1(t_k)\Omega_0^{-2}$, and used the fact that the expressions (IV.5.17–IV.5.18) have only zero- and first-order terms. The first order term in (IV.C.3) is thus

$$\sum_{k=1}^{n}(-\Omega_0^2)^{n-k}[\mu_1(t_k) - m_1(\tilde{t}_k)](-\Omega_0^2)^k, \tag{IV.C.4}$$

which gives

$$\mathfrak{B}_1 = \sum_{n=1}^{\infty}\left(\prod_{k=1}^{n}\int_{t_a}^{\tilde{t}_{k-1}} dt_k \int_{t_a}^{t_k} d\tilde{t}_k\right)\sum_{k=1}^{n}(-\Omega_0^2)^{n-k}[\mu_1(t_k) - m_1(\tilde{t}_k)](-\Omega_0^2)^k. \tag{IV.C.5}$$

If we use[12]

$$\begin{aligned}\left(\prod_{k=1}^{n}\int_{t_a}^{\tilde t_{k-1}} d\tilde t_k\right) f(\tilde t_\ell) &= \left(\prod_{k=1}^{\ell}\int_{t_a}^{\tilde t_{k-1}} d\tilde t_k\right) f(\tilde t_\ell)\left(\prod_{k=\ell+1}^{n}\int_{t_a}^{\tilde t_{k-1}} d\tilde t_k\right)\\ &= \left(\prod_{k=1}^{\ell}\int_{t_a}^{\tilde t_{k-1}} d\tilde t_k\right) f(\tilde t_\ell)\frac{(\tilde t_\ell - t_a)^{n-\ell}}{(n-\ell)!} = \left(\prod_{k=1}^{\ell}\int_{\hat t_{k-1}}^{t_b} d\hat t_k\right) f(\hat t_1)\frac{(\hat t_1 - t_a)^{n-\ell}}{(n-\ell)!}\\ &= \int_{t_a}^{t_b} d\hat t_1 f(\hat t_1)\frac{(\hat t_1 - t_a)^{n-\ell}}{(n-\ell)!}\left(\prod_{k=2}^{\ell}\int_{\hat t_{k-1}}^{t_b} d\hat t_k\right) = \int_{t_a}^{t_b} dt f(t)\frac{(t_b - t)^{\ell-1}}{(\ell-1)!}\frac{(t-t_a)^{n-\ell}}{(n-\ell)!},\end{aligned} \tag{IV.C.6}$$

we find that

$$\begin{aligned}\mathfrak{B}_1 &= \sum_{n=1}^{\infty}\frac{(-1)^n}{(2n-1)!}\int_{t_a}^{t_b} dt\sum_{k=1}^{n}\Big[\Omega_0^{2n-2k}(t-t_a)^{2n-2k+1}\tbinom{2n-1}{2k-2}\mu_1(t)(t_b-t)^{2k-2}\Omega_0^{2k}\\ &\qquad - \Omega_0^{2n-2k}(t-t_a)^{2n-2k}\tbinom{2n-1}{2k-1}m_1(t)(t_b-t)^{2k-1}\Omega_0^{2k}\Big]\\ &= \sum_{n=1}^{\infty}\frac{(-1)^n}{(2n-1)!}\int_{t_a}^{t_b} dt\sum_{k=1}^{n}\Big\{\Omega_0^{-1}[\Omega_0(t-t_a)]^{2n-2k+1}\tbinom{2n-1}{2k-2}\varpi_1(t)[\Omega_0(t_b-t)]^{2k-2}\\ &\qquad - [\Omega_0(t-t_a)]^{2n-2k}\tbinom{2n-1}{2k-1}m_1(t)[\Omega_0(t_b-t)]^{2k-1}\Omega_0\Big\}.\end{aligned} \tag{IV.C.7}$$

To proceed further, we should streamline the notation for components of matrices on $\mathbb{R}^{\mathcal{S}/2\oplus\mathcal{S}/2}$. At the moment, the components of Ω_0 are written as $\Omega^{\mathsf{UL}}_{0\mathbf{M}_1\mathbf{M}_2} = \delta_{\mathbf{M}_1\mathbf{M}_2}k_{\mathbf{M}_1}$, $\Omega^{\mathsf{UR}}_{0\mathbf{M}_1\mathbf{M}_2} = 0$, $\Omega^{\mathsf{LL}}_{0\mathbf{M}_1\mathbf{M}_2} = 0$, and $\Omega^{\mathsf{LR}}_{0\mathbf{M}_1\mathbf{M}_2} = \delta_{\mathbf{M}_1\mathbf{M}_2}k_{\mathbf{M}_1}$. This is greatly simplified if we observe that $k_{-\mathbf{M}} = |-\mathbf{k}_{\mathbf{M}}| = k_{\mathbf{M}}$, and define

$$M = \begin{pmatrix}\{M^{\mathsf{UL}}_{\mathbf{M}_1\mathbf{M}_2}\} & \{M^{\mathsf{UR}}_{\mathbf{M}_1\mathbf{M}_2}\}\\ \{M^{\mathsf{LL}}_{\mathbf{M}_1\mathbf{M}_2}\} & \{M^{\mathsf{LR}}_{\mathbf{M}_1\mathbf{M}_2}\}\end{pmatrix} = \begin{pmatrix}\{M_{\mathbf{M}_1\mathbf{M}_2}\} & \{M_{\mathbf{M}_1,-\mathbf{M}_2}\}\\ \{M_{-\mathbf{M}_1,\mathbf{M}_2}\} & \{M_{-\mathbf{M}_1,-\mathbf{M}_2}\}\end{pmatrix} = \{M_{\mathbf{M}_1\mathbf{M}_2}\}, \tag{IV.C.8}$$

where the indices in the first two expressions range over $\mathcal{S}/2$ and those in the third range

[12] As defined in Appendix IV.B, $\tilde t_0 = t_b$. In addition, $\hat t_0 = t_a$.

over $\mathcal{S}$. Then $\Omega_{0\mathbf{M}_1\mathbf{M}_2} = \delta_{\mathbf{M}_1\mathbf{M}_2} k_{\mathbf{M}_1}$, and

$$\begin{aligned}
&\mathfrak{B}_{1\mathbf{M}_1\mathbf{M}_2} \\
&= \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n-1)!} \int_{t_a}^{t_b} dt \Bigg\{ \frac{\varpi_1(t)_{\mathbf{M}_1\mathbf{M}_2}}{k_{\mathbf{M}_1}} \sum_{k=1}^{n} \tbinom{2n-1}{2k-2} [\overbrace{k_{\mathbf{M}_1}(t-t_a)}^{\theta_{A1}}]^{2n-2k+1} [\overbrace{k_{\mathbf{M}_2}(t_b-t)}^{\theta_{B2}}]^{2k-2} \\
&\qquad\qquad - k_{\mathbf{M}_2} m_1(t)_{\mathbf{M}_1\mathbf{M}_2} \sum_{k=1}^{n} \tbinom{2n-1}{2k-1} [k_{\mathbf{M}_1}(t-t_a)]^{2n-2k} [k_{\mathbf{M}_2}(t_b-t)]^{2k-1} \Bigg\} \\
&= \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n-1)!} \int_{t_a}^{t_b} dt \Bigg(\varpi_1(t)_{\mathbf{M}_1\mathbf{M}_2} k_{\mathbf{M}_1}^{-1} \frac{(\theta_{A1}+\theta_{B2})^{2n-1} + (\theta_{A1}-\theta_{B2})^{2n-1}}{2} \\
&\qquad\qquad - m_1(t)_{\mathbf{M}_1\mathbf{M}_2} k_{\mathbf{M}_2} \frac{(\theta_{A1}+\theta_{B2})^{2n-1} - (\theta_{A1}-\theta_{B2})^{2n-1}}{2} \Bigg) \\
&= -\int_{t_a}^{t_b} dt \Bigg(\varpi_1(t)_{\mathbf{M}_1\mathbf{M}_2} k_{\mathbf{M}_1}^{-1} \frac{\sin(\theta_{A1}+\theta_{B2}) + \sin(\theta_{A1}-\theta_{B2})}{2} \\
&\qquad\qquad - m_1(t)_{\mathbf{M}_1\mathbf{M}_2} k_{\mathbf{M}_2} \frac{\sin(\theta_{A1}+\theta_{B2}) - \sin(\theta_{A1}-\theta_{B2})}{2} \Bigg) \\
&= \int_{t_a}^{t_b} dt \left[m_1(t)_{\mathbf{M}_1\mathbf{M}_2} k_{\mathbf{M}_2} \cos\theta_{A1} \sin\theta_{B2} - \varpi_1(t)_{\mathbf{M}_1\mathbf{M}_2} k_{\mathbf{M}_1}^{-1} \sin\theta_{A1} \cos\theta_{B2} \right]
\end{aligned} \tag{IV.C.9}$$

Then we can use the first order term in (IV.B.19b)

$$\mathfrak{C}_1 = -\frac{d\mathfrak{B}_1}{dt_b} \Omega_0^{-2} - \Omega_0 \sin\Omega_0 T_{ba} \Omega_0^{-2} \varpi_1(t_b) \Omega_0^{-2} \tag{IV.C.10}$$

to calculate

$$\mathfrak{C}_{1\mathbf{M}_1\mathbf{M}_2} = -\int_{t_a}^{t_b} dt \left[m_1(t)_{\mathbf{M}_1\mathbf{M}_2} \cos\theta_{A1} \cos\theta_{B2} + \varpi_1(t)_{\mathbf{M}_1\mathbf{M}_2} \frac{\sin\theta_{A1}\sin\theta_{B2}}{k_{\mathbf{M}_1} k_{\mathbf{M}_2}} \right]; \tag{IV.C.11}$$

likewise, the first order term in (IV.B.19c),

$$\mathfrak{A}_1 = -\frac{d\mathfrak{C}_1}{dt_a} + m_1(t_a) \cos\Omega_0 T_{ba}, \tag{IV.C.12}$$

gives

$$\mathfrak{A}_{1\mathbf{M}_1\mathbf{M}_2} = \int_{t_a}^{t_b} dt \left[m_1(t)_{\mathbf{M}_1\mathbf{M}_2} k_{\mathbf{M}_1} \sin\theta_{A1} \cos\theta_{B2} - \varpi_1(t)_{\mathbf{M}_1\mathbf{M}_2} k_{\mathbf{M}_2}^{-1} \cos\theta_{A1} \sin\theta_{B2} \right]. \tag{IV.C.13}$$

It is straightforward to check that (IV.C.9), (IV.C.11), and (IV.C.13) satisfy (IV.B.13) and (IV.B.14)

IV.C.2 A, B and C

We can use the expansions for $\mathfrak{A}$, $\mathfrak{B}$ and $\mathfrak{C}$ calculated in Section IV.C.1 to find expansions for

$$A[\Phi] = \mathfrak{A}\left(\frac{T}{2}\middle|-\frac{T}{2}\right)\mathfrak{C}^{-1}\left(\frac{T}{2}\middle|-\frac{T}{2}\right) - \dot{m}\left(-\frac{T}{2}\right) \tag{IV.5.25a}$$

$$B[\Phi] = \mathfrak{C}^{-1}\left(\frac{T}{2}\middle|-\frac{T}{2}\right)\mathfrak{B}\left(\frac{T}{2}\middle|-\frac{T}{2}\right) + \dot{m}\left(\frac{T}{2}\right) \tag{IV.5.25b}$$

and

$$C[\Phi] = \mathfrak{C}^{-1}\left(\frac{T}{2}\middle|-\frac{T}{2}\right). \tag{IV.5.25c}$$

From (IV.C.2), the zero order terms are

$$A_0 = \frac{\Omega_0}{\tan\Omega_0 T} = B_0 \tag{IV.C.14a}$$

$$C_0 = \frac{\Omega_0}{\sin\Omega_0 T}; \tag{IV.C.14b}$$

Proceeding to the first order terms, we have

$$A_1[\Phi] = \mathfrak{A}_1\frac{\Omega_0}{\sin\Omega_0 T} - \cos\Omega_0 T\frac{\Omega_0}{\sin\Omega_0 T}\mathfrak{C}_1\frac{\Omega_0}{\sin\Omega_0 T} - \dot{m}_1\left(-\frac{T}{2}\right); \tag{IV.C.15}$$

using (IV.C.13) and (IV.C.11) gives

$$\begin{aligned} A_{1\mathbf{M}_1\mathbf{M}_2} = \frac{k_{\mathbf{M}_1}k_{\mathbf{M}_2}}{\sin k_{\mathbf{M}_1}T\sin k_{\mathbf{M}_2}T}\int_{-T/2}^{T/2} dt\Big[& m_1(t)_{\mathbf{M}_1\mathbf{M}_2}\sin k_{\mathbf{M}_1}T\sin\theta_{A1}\cos\theta_{B2} \\ -\frac{\varpi_1(t)_{\mathbf{M}_1\mathbf{M}_2}}{k_{\mathbf{M}_1}k_{\mathbf{M}_2}}\sin k_{\mathbf{M}_1}T\cos\theta_{A1}\sin\theta_{B2} & + m_1(t)_{\mathbf{M}_1\mathbf{M}_2}\cos k_{\mathbf{M}_1}T\cos\theta_{A1}\cos\theta_{B2} \\ & + \frac{\varpi_1(t)_{\mathbf{M}_1\mathbf{M}_2}}{k_{\mathbf{M}_1}k_{\mathbf{M}_2}}\cos k_{\mathbf{M}_1}T\sin\theta_{A1}\sin\theta_{B2}\Big] - \dot{m}_1\left(-\frac{T}{2}\right)_{\mathbf{M}_1\mathbf{M}_2}; \end{aligned} \tag{IV.C.16}$$

if we note that $k_{\mathbf{M}_1}T - \theta_{A1} = k_{\mathbf{M}_1}(T/2 - t) = \theta_{B1}$, we can convert this to

$$\begin{aligned} A_{1\mathbf{M}_1\mathbf{M}_2} = \frac{k_{\mathbf{M}_1}k_{\mathbf{M}_2}}{\sin k_{\mathbf{M}_1}T\sin k_{\mathbf{M}_2}T}\int_{-T/2}^{T/2} dt\Big[& m_1(t)_{\mathbf{M}_1\mathbf{M}_2}\cos\theta_{B1}\cos\theta_{B2} \\ & - \frac{\varpi_1(t)_{\mathbf{M}_1\mathbf{M}_2}}{k_{\mathbf{M}_1}k_{\mathbf{M}_2}}\sin\theta_{B1}\sin\theta_{B2}\Big] - \dot{m}_1\left(-\frac{T}{2}\right)_{\mathbf{M}_1\mathbf{M}_2}. \end{aligned} \tag{IV.C.17}$$

The $\dot{m}$ boundary term can be converted as follows:

$$
\begin{aligned}
&-\dot{m}_1\left(-\frac{T}{2}\right)_{\mathbf{M}_1\mathbf{M}_2}\sin k_{\mathbf{M}_1}T\sin k_{\mathbf{M}_2}T=\dot{m}(t)_{\mathbf{M}_1\mathbf{M}_2}\sin\theta_{B1}\sin\theta_{B2}\big|_{-T/2}^{T/2}\\
&\quad=\int_{-T/2}^{T/2}dt\left[\ddot{m}_1(t)_{\mathbf{M}_1\mathbf{M}_2}\sin\theta_{B1}\sin\theta_{B2}+\dot{m}_1(t)_{\mathbf{M}_1\mathbf{M}_2}\frac{d}{dt}\left(\sin\theta_{B1}\sin\theta_{B2}\right)\right];
\end{aligned}
\tag{IV.C.18}
$$

integration by parts gives

$$
\begin{aligned}
&\int_{-T/2}^{T/2}dt\dot{m}_1(t)_{\mathbf{M}_1\mathbf{M}_2}\frac{d}{dt}\left(\sin\theta_{B1}\sin\theta_{B2}\right)\\
&=m_1(t)_{\mathbf{M}_1\mathbf{M}_2}\frac{d}{dt}\left(\sin\theta_{B1}\sin\theta_{B2}\right)\Big|_{-T/2}^{T/2}-\int_{-T/2}^{T/2}dtm_1(t)_{\mathbf{M}_1\mathbf{M}_2}\frac{d^2}{dt^2}\left(\sin\theta_{B1}\sin\theta_{B2}\right)\\
&=m_1\left(-\frac{T}{2}\right)_{\mathbf{M}_1\mathbf{M}_2}\left(k_{\mathbf{M}_1}\cos k_{\mathbf{M}_1}T\sin k_{\mathbf{M}_2}T+k_{\mathbf{M}_2}\sin k_{\mathbf{M}_1}T\cos k_{\mathbf{M}_2}T\right)\\
&+\int_{-T/2}^{T/2}dtm_1(t)_{\mathbf{M}_1\mathbf{M}_2}\left[(k_{\mathbf{M}_1}^2+k_{\mathbf{M}_2}^2)\sin\theta_{B1}\sin\theta_{B2}-2k_{\mathbf{M}_1}k_{\mathbf{M}_2}\cos\theta_{B1}\cos\theta_{B2}\right],
\end{aligned}
\tag{IV.C.19}
$$

so

$$
\begin{aligned}
&A_{1\mathbf{M}_1\mathbf{M}_2}=m_1\left(-\frac{T}{2}\right)_{\mathbf{M}_1\mathbf{M}_2}\left(\frac{k_{\mathbf{M}_1}}{\tan k_{\mathbf{M}_1}T}+\frac{k_{\mathbf{M}_2}}{\tan k_{\mathbf{M}_2}T}\right)\\
&-\frac{k_{\mathbf{M}_1}k_{\mathbf{M}_2}}{\sin k_{\mathbf{M}_1}T\sin k_{\mathbf{M}_2}T}\int_{-T/2}^{T/2}dt\Big[m_1(t)_{\mathbf{M}_1\mathbf{M}_2}\cos\theta_{B1}\cos\theta_{B2}+n_1(t)_{\mathbf{M}_1\mathbf{M}_2}\sin\theta_{B1}\sin\theta_{B2}\Big]
\end{aligned}
\tag{IV.C.20}
$$

where

$$
n_1(t)_{\mathbf{M}_1\mathbf{M}_2}=\frac{\varpi_1(t)_{\mathbf{M}_1\mathbf{M}_2}-\ddot{m}_1(t)_{\mathbf{M}_1\mathbf{M}_2}-m_1(t)_{\mathbf{M}_1\mathbf{M}_2}(k_{\mathbf{M}_1}^2+k_{\mathbf{M}_2}^2)}{k_{\mathbf{M}_1}k_{\mathbf{M}_2}}.
\tag{IV.C.21}
$$

Analogously, we find

$$
\begin{aligned}
&B_{1\mathbf{M}_1\mathbf{M}_2}=m_1\left(\frac{T}{2}\right)_{\mathbf{M}_1\mathbf{M}_2}\left(\frac{k_{\mathbf{M}_1}}{\tan k_{\mathbf{M}_1}T}+\frac{k_{\mathbf{M}_2}}{\tan k_{\mathbf{M}_2}T}\right)\\
&-\frac{k_{\mathbf{M}_1}k_{\mathbf{M}_2}}{\sin k_{\mathbf{M}_1}T\sin k_{\mathbf{M}_2}T}\int_{-T/2}^{T/2}dt\Big[m_1(t)_{\mathbf{M}_1\mathbf{M}_2}\cos\theta_{A1}\cos\theta_{A2}+n_1(t)_{\mathbf{M}_1\mathbf{M}_2}\sin\theta_{A1}\sin\theta_{A2}\Big]
\end{aligned}
\tag{IV.C.22}
$$

The first order term $C_1 = -\frac{\Omega_0}{\sin\Omega_0 T}\mathfrak{C}_1\frac{\Omega_0}{\sin\Omega_0 T}$ can be cast into a similar form by noticing that

$$\begin{aligned}
&\int_{-T/2}^{T/2} dt\ddot{m}_1(t)_{\mathbf{M}_1\mathbf{M}_2}\sin\theta_{A1}\sin\theta_{B2}\\
&= \dot{m}_1(t)_{\mathbf{M}_1\mathbf{M}_2}\sin\theta_{A1}\sin\theta_{B2}\big|_{-T/2}^{T/2} - \int_{-T/2}^{T/2} dt\dot{m}_1(t)_{\mathbf{M}_1\mathbf{M}_2}\frac{d}{dt}\left(\sin\theta_{A1}\sin\theta_{B2}\right)\\
&= -m_1(t)_{\mathbf{M}_1\mathbf{M}_2}\left(k_{\mathbf{M}_1}\cos\theta_{A1}\sin\theta_{B2} - k_{\mathbf{M}_2}\sin\theta_{A1}\cos\theta_{B2}\right)\big|_{-T/2}^{T/2}\\
&\qquad\qquad + \int_{-T/2}^{T/2} dt m_1(t)_{\mathbf{M}_1\mathbf{M}_2}\frac{d^2}{dt^2}\left(\sin\theta_{A1}\sin\theta_{B2}\right)\\
&= m_1\left(\frac{T}{2}\right)_{\mathbf{M}_1\mathbf{M}_2} k_{\mathbf{M}_2}\sin k_{\mathbf{M}_1}T + m_1\left(-\frac{T}{2}\right)_{\mathbf{M}_1\mathbf{M}_2} k_{\mathbf{M}_1}\sin k_{\mathbf{M}_2}T\\
&- \int_{-T/2}^{T/2} dt m_1(t)_{\mathbf{M}_1\mathbf{M}_2}\left[(k_{\mathbf{M}_1}^2 + k_{\mathbf{M}_2}^2)\sin\theta_{A1}\sin\theta_{B2} + 2k_{\mathbf{M}_1}k_{\mathbf{M}_2}\cos\theta_{A1}\cos\theta_{B2}\right],
\end{aligned}\tag{IV.C.23}$$

so that

$$\begin{aligned}
&C_{1\mathbf{M}_1\mathbf{M}_2} = m_1\left(\frac{T}{2}\right)_{\mathbf{M}_1\mathbf{M}_2}\frac{k_{\mathbf{M}_2}}{\sin k_{\mathbf{M}_2}T} + m_1\left(-\frac{T}{2}\right)_{\mathbf{M}_1\mathbf{M}_2}\frac{k_{\mathbf{M}_1}}{\sin k_{\mathbf{M}_1}T}\\
&- \frac{k_{\mathbf{M}_1}k_{\mathbf{M}_2}}{\sin k_{\mathbf{M}_1}T\sin k_{\mathbf{M}_2}T}\int_{-T/2}^{T/2} dt\Big[m_1(t)_{\mathbf{M}_1\mathbf{M}_2}\cos\theta_{A1}\cos\theta_{B2} - n_1(t)_{\mathbf{M}_1\mathbf{M}_2}\sin\theta_{A1}\sin\theta_{B2}\Big].
\end{aligned}\tag{IV.C.24}$$

Now, each of the expressions (IV.C.20), (IV.C.22) and (IV.C.24) contains integrals over terms with two t-dependent trigonometric functions, *e.g.*, $\cos\theta_{A1}\cos\theta_{B2} = \cos k_{\mathbf{M}_1}(t + T/2)\cos k_{\mathbf{M}_2}(T/2 - t)$; this can be rewritten using

$$\begin{aligned}
&\cos k_{\mathbf{M}_1}(t + T/2)\cos k_{\mathbf{M}_2}(T/2 - t) = \frac{1}{2}\left[\cos(k_- t + k_+ T/2) + \cos(k_+ t + k_- T/2)\right]\\
&\quad = \frac{1}{2}\left[\cos k_- t\cos\frac{k_+ T}{2} - \sin k_- t\sin\frac{k_+ T}{2} + \cos k_+ t\cos\frac{k_- T}{2} - \cos k_+ t\cos\frac{k_- T}{2}\right],
\end{aligned}\tag{IV.C.25}$$

where $k_\pm = k_{\mathbf{M}_1} \pm k_{\mathbf{M}_2}$. This sort of identity allows us to rewrite (IV.C.20), (IV.C.22) and (IV.C.24) as

$$\begin{aligned}
&A_{1\mathbf{M}_1\mathbf{M}_2} = m_1\left(-\frac{T}{2}\right)_{\mathbf{M}_1\mathbf{M}_2}\left(\frac{k_{\mathbf{M}_1}}{\tan k_{\mathbf{M}_1}T} + \frac{k_{\mathbf{M}_2}}{\tan k_{\mathbf{M}_2}T}\right)\\
&- \frac{k_{\mathbf{M}_1}k_{\mathbf{M}_2}}{2\sin k_{\mathbf{M}_1}T\sin k_{\mathbf{M}_2}T}\int_{-T/2}^{T/2} dt\Big\{\left[m_1(t) + n_1(t)\right]_{\mathbf{M}_1\mathbf{M}_2}\left(\cos k_- t\cos\frac{k_- T}{2} + \sin k_- t\sin\frac{k_- T}{2}\right)\\
&\qquad\qquad + \left[m_1(t) - n_1(t)\right]_{\mathbf{M}_1\mathbf{M}_2}\left(\cos k_+ t\cos\frac{k_+ T}{2} + \sin k_+ t\sin\frac{k_+ T}{2}\right)\Big\}
\end{aligned}\tag{IV.C.26a}$$

$$
\begin{aligned}
B_{1\mathbf{M}_1\mathbf{M}_2} &= m_1\left(\frac{T}{2}\right)_{\mathbf{M}_1\mathbf{M}_2}\left(\frac{k_{\mathbf{M}_1}}{\tan k_{\mathbf{M}_1}T}+\frac{k_{\mathbf{M}_2}}{\tan k_{\mathbf{M}_2}T}\right)\\
-&\frac{k_{\mathbf{M}_1}k_{\mathbf{M}_2}}{2\sin k_{\mathbf{M}_1}T\sin k_{\mathbf{M}_2}T}\int_{-T/2}^{T/2}dt\Big\{[m_1(t)+n_1(t)]_{\mathbf{M}_1\mathbf{M}_2}\left(\cos k_-t\cos\frac{k_-T}{2}-\sin k_-t\sin\frac{k_-T}{2}\right)\\
&\qquad+[m_1(t)-n_1(t)]_{\mathbf{M}_1\mathbf{M}_2}\left(\cos k_+t\cos\frac{k_+T}{2}-\sin k_+t\sin\frac{k_+T}{2}\right)\Big\}
\end{aligned}
\tag{IV.C.26b}
$$

$$
\begin{aligned}
C_{1\mathbf{M}_1\mathbf{M}_2} &= m_1\left(\frac{T}{2}\right)_{\mathbf{M}_1\mathbf{M}_2}\frac{k_{\mathbf{M}_2}}{\sin k_{\mathbf{M}_2}T}+m_1\left(-\frac{T}{2}\right)_{\mathbf{M}_1\mathbf{M}_2}\frac{k_{\mathbf{M}_1}}{\sin k_{\mathbf{M}_1}T}\\
-&\frac{k_{\mathbf{M}_1}k_{\mathbf{M}_2}}{2\sin k_{\mathbf{M}_1}T\sin k_{\mathbf{M}_2}T}\int_{-T/2}^{T/2}dt\Big\{[m_1(t)+n_1(t)]_{\mathbf{M}_1\mathbf{M}_2}\left(\cos k_-t\cos\frac{k_+T}{2}-\sin k_-t\sin\frac{k_+T}{2}\right)\\
&\qquad+[m_1(t)-n_1(t)]_{\mathbf{M}_1\mathbf{M}_2}\left(\cos k_+t\cos\frac{k_-T}{2}-\sin k_+t\sin\frac{k_-T}{2}\right)\Big\}
\end{aligned}
\tag{IV.C.26c}
$$

These seem to be taking on a nice form in terms of more basically defined modes

$$
\chi_\pm=\int_{-T/2}^{T/2}dt[m_1(t)\mp n_1(t)]_{\mathbf{M}_1\mathbf{M}_2}\cos k_\pm t+\text{boundary terms}
\tag{IV.C.27a}
$$

and

$$
\sigma_\pm=\int_{-T/2}^{T/2}dt[m_1(t)\mp n_1(t)]_{\mathbf{M}_1\mathbf{M}_2}\sin k_\pm t+\text{boundary terms},
\tag{IV.C.27b}
$$

namely

$$
\begin{pmatrix}B_{1\mathbf{M}_1\mathbf{M}_2}\\ A_{1\mathbf{M}_1\mathbf{M}_2}\\ C_{1\mathbf{M}_1\mathbf{M}_2}\\ C_{1\mathbf{M}_1\mathbf{M}_2}^{\text{tr}}\end{pmatrix}=-\frac{k_{\mathbf{M}_1}k_{\mathbf{M}_2}}{2\sin k_{\mathbf{M}_1}T\sin k_{\mathbf{M}_2}T}\begin{pmatrix}\cos\frac{k_-T}{2} & -\sin\frac{k_-T}{2} & \cos\frac{k_+T}{2} & -\sin\frac{k_+T}{2}\\ \cos\frac{k_-T}{2} & \sin\frac{k_-T}{2} & \cos\frac{k_+T}{2} & \sin\frac{k_+T}{2}\\ \cos\frac{k_+T}{2} & -\sin\frac{k_+T}{2} & \cos\frac{k_-T}{2} & -\sin\frac{k_-T}{2}\\ \cos\frac{k_+T}{2} & \sin\frac{k_+T}{2} & \cos\frac{k_-T}{2} & \sin\frac{k_-T}{2}\end{pmatrix}\begin{pmatrix}\chi_-\\ \sigma_-\\ \chi_+\\ \sigma_+\end{pmatrix},
\tag{IV.C.28}
$$

Where we calculate $C_{1\mathbf{M}_1\mathbf{M}_2}^{\text{tr}}=C_{1\mathbf{M}_2\mathbf{M}_1}$. We can use this to define $\chi_\pm$ and $\sigma_\pm$, and inverting the matrix to write

$$
\begin{pmatrix}\chi_-\\ \sigma_-\\ \chi_+\\ \sigma_+\end{pmatrix}=-\frac{1}{2k_{\mathbf{M}_1}k_{\mathbf{M}_2}}\begin{pmatrix}\cos\frac{k_-T}{2} & \cos\frac{k_-T}{2} & -\cos\frac{k_+T}{2} & -\cos\frac{k_+T}{2}\\ \sin\frac{k_-T}{2} & -\sin\frac{k_-T}{2} & -\sin\frac{k_+T}{2} & \sin\frac{k_+T}{2}\\ -\cos\frac{k_+T}{2} & -\cos\frac{k_+T}{2} & \cos\frac{k_-T}{2} & \cos\frac{k_-T}{2}\\ -\sin\frac{k_+T}{2} & \sin\frac{k_+T}{2} & \sin\frac{k_-T}{2} & -\sin\frac{k_-T}{2}\end{pmatrix}\begin{pmatrix}B_{1\mathbf{M}_1\mathbf{M}_2}\\ A_{1\mathbf{M}_1\mathbf{M}_2}\\ C_{1\mathbf{M}_1\mathbf{M}_2}\\ C_{1\mathbf{M}_1\mathbf{M}_2}^{\text{tr}}\end{pmatrix}
\tag{IV.C.29}
$$

determines the boundary terms, giving

$$
\chi_\pm=\int_{-T/2}^{T/2}dt[m_1(t)\mp n_1(t)]_{\mathbf{M}_1\mathbf{M}_2}\cos k_\pm t\mp m_1(t)_{\mathbf{M}_1\mathbf{M}_2}\frac{k_\pm}{k_{\mathbf{M}_1}k_{\mathbf{M}_2}}\sin 2k_\pm t\Big|_{-T/2}^{T/2}
\tag{IV.C.30a}
$$

and

$$\sigma_{\pm} = \int_{-T/2}^{T/2} dt [m_1(t) \mp n_1(t)]_{\mathbf{M}_1\mathbf{M}_2} \sin k_{\pm} t \pm m_1(t)_{\mathbf{M}_1\mathbf{M}_2} \frac{k_{\pm}}{k_{\mathbf{M}_1} k_{\mathbf{M}_2}} \cos 2k_{\pm} t \Big|_{-T/2}^{T/2} . \tag{IV.C.30b}$$

Chapter V

What have we learned?

In conclusion, these are some lessons and accomplishments of this research program.

The first project (Chapter II) concerned spacetime alternatives in the generalized quantum mechanics of a relativistic particle. While the theory was chosen for its reparametrization invariance, in analogy to the time reparametrization invariance of general relativity, that invariance was not the focus of the work. Rather, the goal was to exhibit a simple coarse graining which would decohere for certain initial and final conditions. Due to the complexity of the theory (notably the inclusion in the sum of paths which go both forward and backward in time), the decoherence functional could only be calculated for coarse grainings involving a region extending arbitrarily far into the past and future. This leads to a number of further complications, such as a non-trivial final condition which allowed non-overlapping branch wavefunctions to interfere with one another, and a possible dependence of the decoherence functional upon the choice of surfaces on which the initial and final conditions are attached. Nonetheless it is possible to choose initial conditions which lead to decoherence and the assignment of probabilities.

When generalized quantum mechanics was applied to a non-Abelian gauge theory (Chapter III), the gauge invariance of the theory, and the constraints associated therewith, were of paramount interest. First (Sec III.3), I demonstrated one way to convert the formal sum-over-histories expression for the class operators used to construct the decoherence functional to a more explicit "skeletonized" construction on a spacetime lattice. In the limit that the lattice spacing vanishes, this construction is gauge invariant and gives the same result for the propagator as a reduced phase space operator theory. Then I turned to results for particular coarse grainings, concentrating on those defined by the constraints. When the constraints are expressed in terms of the full phase space variables (Sec III.4), with the gauge electric field represented by the momentum conjugate to the vector potential, the scalar potential acts as a Lagrange multiplier, enforcing the constraint identically. This holds for any coarse graining which does not involve a covariant time derivative. If the constraints are defined solely in terms of configuration space variables (Sec. III.5), however, with the gauge electric field expressed in terms of the scalar and vector potentials, things are less clear-cut. For coarse grainings defined by the non-Abelian equivalent of the longitudinal part of the electric field, I extended the result from Abelian theory, namely that any coarse graining which decoheres produces a definite prediction that the constraint is satisfied, and exhibited such a decohering coarse graining. However, the result did not extend to the non-Abelian *divergence* of the electric field because of the covariant derivative's dependence on what are usually called the physical degrees of freedom. And in the Abelian case, I constructed a quantity which should vanish in the presence of the constraints, but which also involves the "physical" degrees of freedom, and demonstrated that there were decohering coarse grainings by this quantity which predicted non-zero probabilities for alternatives inconsistent with the constraint. This is a different result than would be predicted for the analogous phase space quantity. Insight into the nature of this dilemma was gained by demonstrating the manifest Lorentz invariance (Sec III.6) of the sum-over-histories generalized quantum mechanics of this theory when expressed in the solely configuration-space form. Even the operator constraints imposed upon the initial and final wavefunctions do not pick out an arbitrary time direction. Instead, the relevant direction is the local normal to the spacelike surface on which the wavefunctions are defined. (These results were demonstrated only formally, and not in a skeletonized theory.) I observed that, from the point of view of Lorentz invariance, it is not surprising that the constraints, expressed as configuration space quantities, are not always satisfied, as there is nothing to distinguish them from any of the other components of the equations of motion. Since defining alternatives at a moment of time would select a preferred time direction, it is reasonable to conjecture that the spacetime nature of the field averages is an essential part of this result.

The final project, concerning the decoherence of spacetime (Chapter IV), demonstrated a toy model result which suggested that coarse grainings by long-wavelength features of the gravitational field might be made to decohere via their coupling to the unobserved short-wavelength modes. The toy model used has an action similar to that for perturbative general relativity, with the perturbed metric replaced by a scalar field. Due to the perturbative nature of the calculation, the non-perturbative effect of decoherence can only occur if some other quantity, such as the inverse temperature of the short-wavelength "environment", becomes small. If the short-wavelength modes are taken to be in a thermal state defined by the expected cosmic graviton background temperature, the scale dividing "long" and "short" wavelengths must be larger than around a millimeter. In that case, sufficiently coarse-grained sets of alternatives involving long-wavelength modes whose spatial frequency is higher than their temporal frequency (which includes static configurations) should decohere.

So I have illustrated some of the capabilities of, and challenges to, generalized quantum mechanics as a quantization method for a theory of gravity, in particular in the areas of spacetime coarse grainings, the role of the constraints, and the use of some gravitational modes to induce decoherence in others. Of course, many issues remain. The relativistic particle could be formulated in a manner which more heavily stressed the reparametrization invariance, the relationship between spacetime alternatives in a gauge theory and the singling out of the constraint could be further explored, and there remains much ground to cover on the decoherence of spacetime, notably to restore the tensor nature of the field, and eventually to attempt a non-perturbative analysis. In addition, there are many open questions in the quantization of gravity and similar theories which were not addressed here, including the resolution of the non-renormalizability problem and the impact of global topological effects.

Bibliography

[1] J. B. Hartle, Phys. Rev. D **44,** 3173 (1991).

[2] N. Yamada and S. Takagi, Prog. Theor. Phys. **87,** 77 (1992). See also N. Yamada and S. Takagi, Prog. Theor. Phys. **85,** 985 (1991); **86,** 599 (1991).

[3] R. Omnès, *The Interpretation of Quantum Mechanics* (Princeton University Press, Princeton, 1994)

[4] R. P. Feynman, Rev. Mod. Phys. **20,** 367 (1948).

[5] See, for example, J. B. Hartle, in *Quantum Cosmology and Baby Universes,* proceedings of the 7th Jerusalem Winter School, 1989, edited by S. Coleman, J. Hartle, T. Piran, and S. Weinberg (World Scientific, Singapore, 1991).

[6] R. Griffiths, J. Stat. Phys. **36,** 219 (1984).

[7] R. Omnès, J. Stat. Phys. **53,** 893 (1988); **53,** 933 (1988); **53,** 957 (1988); **57,** 357 (1989); Rev. Mod. Phys. **64,** 339 (1992).

[8] M. Gell-Mann and J. B. Hartle, in *Complexity, Entropy and the Physics of Information,* edited by W. Żurek, SFI Studies in the Sciences of Complexity, Vol. VIII (Addison-Wesley, Reading, MA); in *Proceedings of the 3rd International Symposium on the Foundations of Quantum Mechanics in the Light of New Technology,* edited by S. Kobayashi, H. Ezawa, Y. Murayama, and S. Nomura (Physical Society of Japan, Tokyo, 1990)

[9] W. Żurek, Phys. Rev. D **14,** 1516 (1981); **26,** 1862 (1982).

[10] F. Dowker and A. Kent, Phys. Rev. Lett. **75,** 3038 (1995); J. Stat. Phys. J. Stat. Phys. **82,** 1575 (1996).

[11] M. Gell-Mann and J. B. Hartle, "Strong Decoherence", preprint gr-qc/9509054.

[12] M. Gell-Mann and J. B. Hartle, Phys. Rev. D **47,** 3345 (1993).

[13] M. Gell-Mann and J. B. Hartle, in *Proceedings of the NATO Workshop on the Physical Origin of Time Asymmetry, Mazagon, Spain, 30 Sep - 4 Oct 1991,* edited by J.J. Halliwell, J. Perez-Mercader, and W.H. Żurek. (Cambridge University Press, Cambridge, England, 1994)

[14] R. M. Wald, *General Relativity* (University of Chicago Press, Chicago, 1984).

[15] F. Halzen and A.D. Martin, *Quarks and Leptons: An Introductory Course in Particle Physics* (Wiley, New York, 1984).

[16] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time* (Cambridge University Press, Cambridge, England, 1973).

[17] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, New York, 1973)

[18] See G. F. Smoot and D. Scott, "Cosmic Background Radiation," in R. M. Barnett *et al.* Phys. Rev. D **54,** 1 (1996) and references therein.

[19] K. Kuchař, in *Quantum Gravity 2*, edited by C. J. Isham, R. Penrose and D. W. Sciama (Clarendon, Oxford, 1981).

[20] See, for example, M.B. Green, J.H. Schwartz and E. Witten, *Superstring Theory* (Cambridge University Press, Cambridge, England, 1987).

[21] A. Ashtekar, *Lectures on Non-Perturbative Canonical Gravity* (World Scientific, Singapore, 1991). For a more elementary overview, see G. T. Horowitz, "Ashtekar's Approach to Quantum Gravity" in *Proceedings of the 1991 Strings and Symmetry Conference*, preprint UCSBTH-91-40 (1991).

[22] For a recent review and bibliography, see C. Rovelli, "Loop Quantum Gravity and Black Hole Physics", preprint gr-qc/9608032 (1996).

[23] S. W. Hawking "The path-integral approach to quantum gravity," in *General Relativity, an Einstein Centenary Survey*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1979).

[24] J. B. Hartle, in *Gravitation and Quantizations,* proceedings of the 1992 Les Houches Summer School, edited by B. Julia and J. Zinn-Justin (North Holland, Amsterdam, 1995).

[25] L. D. Fadeev, Teor. Mat. Fiz. **1,** 3 (1969) [Theor. Math. Phys. **1,** 1 (1970)].

[26] J. J. Halliwell, Phys. Rev. D **39,** 2912 (1989).

[27] T. Padmanabhan, Phys. Rev. D **39,** 2924 (1989).

[28] M. J. Gotay and J. E. Marsden, in *Mathematical Aspects of Classical Field Theory*, M. J. Gotay, J. E. Marsden, and V. Moncrief, eds. [Contemp. Math. **132,** 367 (1992)].

[29] S. Weinberg, Rev. Mod. Phys. **61,** 1 (1989).

[30] M. Aguilar-Benitez, *et al.*, Phys. Rev. D **50,** 1173 (1994).

[31] D. Marolf, "Refined Algebraic Quantization: systems with a single constraint", preprint gr-qc/9508015 (1995).

[32] J. J. Halliwell and M. E. Ortiz, Phys. Rev. D **48,** 748 (1993).

[33] R. Arnowitt, S. Deser, and C. W. Misner, in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962).

[34] R. Jackiw, Rev. Mod. Phys. **52,** 661 (1980).

[35] O. Bertolami and P. V. Moniz, Nucl. Phys. B **439,** 259 (1995).

[36] B. Hatfield, *Quantum Field Theory of Point Particles and Strings* (Addison-Wesley, Redwood City, CA, 1992).

[37] H. Georgi, *Lie Algebras in Particle Physics* (Addison-Wesley, Redwood City, CA, 1982).

[38] V. N. Gribov, Nucl. Phys. B **139,** 1 (1978).

[39] M. S. Swanson, *Path Integrals and Quantum Processes* (Academic Press, San Diego, 1992).

[40] F. A. Berezin, Teor. Mat. Fiz. **6,** 194 (1971) [Theor. Math. Phys. **6,** 141 (1971)]; Trudy Moskovskogo Matematicheskogo Obshchestva **17,** 117 (1967) [Transactions of the Moscow Mathematical Society **17,** 129 (1967)]; M. M. Mizrahi, J. Math. Phys. **16,** 2201 (1975).

[41] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965) Section 7-3.

[42] T. D. Lee, *Particle Physics and Introduction to Field Theory,* (Harwood, Chur, Switzerland, 1988) Section 18.1.

[43] J. B. Hartle, Phys. Rev. D **37,** 2818 (1988).

[44] J. J. Halliwell and J. B. Hartle, Phys. Rev. D **43,** 1170 (1991).

[45] P. A. M. Dirac, *Lectures on Quantum Mechanics* (Belfer Graduate School of Science, Yeshiva University, New York, 1964), Chapters 3 and 4.

[46] J. J. Halliwell, Phys. Rev. D **43,** 2590 (1991).

[47] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipies: The Art of Scientific Computing* (Cambridge University Press, Cambridge, England, 1992) Section 12.1.

[48] R. P. Feynman and J. R. Vernon, Ann. Phys. (N.Y.) **24,** 118 (1963).

[49] T. Regge, Nuovo Cimento **19,** 558 (1961). See also J.B. Hartle, J. Math. Phys. **26,** 804 (1985).

[50] E.W. Kolb and M.S. Turner, *The Early Universe* (Addison-Wesley, Redwood City, CA, 1990).