

TESE DE
DOUTORADO

Detecção e detectabilidade em topologia cósmica

BRUNO COELHO CÉSAR MOTA

CENTRO BRASILEIRO DE PESQUISAS FÍSICAS-CBPF
RIO DE JANEIRO, JULHO DE 2007

Dedicatória

*Ao meu avô Cilas, que se foi,
ao Gabriel, que veio,
e à Cecília, que ficou.*

Agradecimentos

Ao Gabriel, à Ceci, a meus pais e irmão, por existirem.

Ao Marcelo, por tudo. O orientador de uma tese é ao mesmo tempo um professor e um conselheiro, e ele soube cumprir ambos os papéis sem (em geral) perder o bom humor.

Ao Reza, pela amizade, e pelo exemplo de vida.

Ao Germán, por ajudar a desbravar o caminho. Nenhum trabalho científico existe em isolamento, e esta tese não é exceção. Provavelmente nenhum texto foi para mim mais didático, e útil para desenvolver minhas próprias idéias, do que a sua tese de doutorado.

Aos meus demais colaboradores: Antônio, Armando, Martin, Maurício e Jailson. Foi uma honra.

Ao Ricardo e à Myriam, pela paciência com meus repetidos atrasos. Em outro contexto, ambos poderiam ser confundidos com reencarnações do Buda.

Ao CNPq, pela bolsa concedida.

À ARPA, e a Sir Tim Berners-Lee e Robert Cailliau, por inventarem a Internet e a World Wide Web.

Aos meus amigos em carne e osso, por me manterem razoavelmente são durante todo o processo, e em particular ao Slava e à Parvin, pelas novas amizades na Córsega.

Ao Nizo, Nobody, Tse, Al Ghaliboon, Pedro, Paula, Alba, Doug, TheIrie, Anônimo, e demais entidades incorpóreas do éter de TCP/IP, por tornarem minha vida mais interessante, e por me ajudarem a ficar acordado nas longas madrugadas.

Resumo

O modelo padrão da cosmologia teve nas últimas décadas grande sucesso em explicar com precisão uma grande quantidade de observações. Mas ainda há algumas discrepâncias entre teoria e observações, cuja explicação talvez seja relativamente simples dentro do paradigma vigente, mas pode também resultar em um entendimento qualitativamente mais profundo do cosmo. O presente trabalho tem portanto início (após uma breve revisão de alguns tópicos fundamentais) com o estudo de uma destas discrepâncias, os aparentes desvios da isotropia estatística na radiação cósmica de fundo. Através da construção de um indicador direcional de anisotropia estatística, mostra-se que o Universo observável possui uma direção preferencial estatisticamente significativa, cuja causa é por enquanto desconhecida.

Se por um lado existem discrepâncias entre observações e teoria, por outro esta tampouco é completa. Saber se o Universo é finito ou infinito, ou, mais especificamente, determinar a topologia de suas seções espaciais, é uma das grandes questões em aberto da cosmologia moderna. Mas ao contrário da geometria, a topologia não é descrita por uma teoria fundamental, e portanto só pode ser deduzida atualmente a partir de observações. A consequência observacional mais imediata da existência de uma topologia não trivial é o surgimento de imagens múltiplas de objetos cósmicos. Existem porém limites intrínsecos de tempo e distância para o alcance de nossas observações astrofísicas, limites estes que podem restringir a detectabilidade da topologia. Desta forma, o trabalho prossegue com o estudo sistemático, em termos de parâmetros de modelos cosmológicos do tipo Λ CDM e gás de Chaplygin generalizado, das condições necessárias para que cada possível variedade ou família de variedades topológicas seja detectável.

É possível ainda estudar o problema inverso da detectabilidade da topologia cósmica. São obtidas restrições numéricas significativas sobre parâmetros dos modelos cosmológicos, ao combinar-se as consequências observacionais de uma hipotética detecção da topologia cósmica com dados advindos da observação de supernovas do tipo SN Ia.

A análise das consequências observacionais de uma topologia cósmica detectável pode ser bastante simplificada se for focada nos aspectos da topologia que efetivamente tem

consequências observáveis, a chamada forma local do Universo. São obtidos alguns resultados gerais que caracterizam a forma local de variedades de curvatura seccional positiva, e é mostrado ainda que, no limite imposto por teorias inflacionárias sobre o valor da curvatura espacial, a forma local das possíveis variedades não-planas pode ser expressa de uma forma simples. Estes resultados permitem aumentar significativamente a eficiência de buscas por evidências de uma topologia cósmica não trivial.

Finalmente, existem diversos efeitos que possivelmente poderiam complicar ou até impedir a detecção da topologia cósmica. O presente trabalho termina com a caracterização completa de um destes efeitos, a aberração dos chamados círculos no céu (um dos principais métodos propostos para a detecção da topologia cósmica) causada pelo movimento próprio de nossa galáxia em relação ao fluxo de Hubble.

Abstract

Whether the Universe is finite or infinite, or, more specifically, determining the topology of its spatial sections, is one of the great open questions in modern cosmology. Unlike geometry, topology is not currently described by a fundamental theory. If on one hand this implies that cosmic topology is at present a purely phenomenological question, on the other hand its eventual detection could possibly be the first direct evidence for some new physics. The most immediate observational consequence of a non-trivial cosmic topology would be the existence of multiple images of cosmic objects. There are, however, intrinsic limits in both time and space to the range of astrophysical observations, and such limits may restrict the detectability of such topology. Broadly speaking, the present work seeks to determine in a systematic fashion, in terms of observational parameters, if, how, and under what conditions, each potential manifold with non-trivial topology is detectable.

Initially, we review some topics which are important to the understanding of cosmic topology in general, and the results that follow in particular. We give particular emphasis on the fundamental concepts of the topology of 3-manifolds of constant curvature, and the connection between the holonomy group of each manifold and its detectability by pattern repetition methods. Next, we present one possible discrepancy between observations and the standard cosmological theory: we construct a direction indicator of statistical anisotropy for the cosmic microwave background radiation. When applied to actual sky maps, this indicator seemingly shows that the observable universe has a statistically-significant preferred direction in which statistical anisotropy is maximum. We do know, as of yet, whether such axis has topological origin or not.

The actual study of detectability then begins in earnest. We present a systematical method for obtaining detectability condition for topologies compatible with a homogeneous, isotropic and non-flat universe, in terms of observable parameters of Λ CDM and generalized Chaplygin gas models. One can also study the converse of this question, and we do so to obtain severe restrictions on cosmological model parameters after combining the observational consequences of the detection a hypothetical cosmic topology with supernova data.

Next, we define the so-called local shape of the universe, which consists of those elements of cosmic topology that are actually detectable by a given observer. We obtain some beautiful results, which enable us to fully characterize the local shape for manifolds with positive space curvature. We further show that in the limit on the absolute curvature of space obtained from typical inflationary models, the local shape of non-flat manifolds take a far more simple form than the description of each manifold as a whole. These results allow for a significant simplification of surveys that seek to detect evidence of a non-trivial cosmic topology.

Finally, there are many different effects that might complicate or even prevent the detection of cosmic topology, even when it is in principle detectable. We fully characterize the consequences of one such effect, our galaxy's proper motion with respect to Hubble's flow, on the so-called circles in the sky, one of the most important methods proposed so far to detect cosmic topology. We show that the actual effect is quite small, but it could, if uncorrected, become relevant in future surveys using upcoming high-precision measurements.

Conteúdo

Dedicatória	i
Agradecimentos	ii
Resumo	iii
Abstract	v
Introdução	1
1 Preliminares	8
1.1 Introdução	8
1.2 Relatividade, e cosmologia relativística	9
1.2.1 Relatividade, beleza e universalidade	9
1.2.2 O Universo como objeto de estudo	11
1.3 A radiação cósmica de fundo	14
1.3.1 A pré-história da RCF	14
1.3.2 A análise das anisotropias	16
1.3.3 Gaussianidade e isotropia estatística	18
1.3.4 A variância cósmica	20
1.4 Topologia, geometria, e topologia cósmica	21
1.4.1 Fundamentos dos fundamentos da topologia	22
1.4.2 Uma história de 2–toros	24
1.4.3 Homotopias por laço, holonomias e espaço de recobrimento	29
1.4.4 Domínio fundamental, e as isometrias da topologia cósmica	35
2 Anisotropia estatística na radiação cósmica de fundo	40
2.1 Introdução	40
2.2 O sucesso da cosmologia, e algumas pequenas discrepâncias	41
2.3 Radiação cósmica de fundo e anisotropia estatística	43
2.4 O σ -mapa	46
2.4.1 O histograma de separação angular de pares	47

2.4.2	O indicador σ	49
2.5	Aplicação aos dados da RCF	51
2.5.1	A escolha de mapas e a contaminação de foreground	51
2.5.2	Calculando os σ -mapas	52
2.5.3	O espectro angular de potência e a relevância estatística dos resultados	53
2.5.4	Escolha de parâmetros e robustez	56
2.6	Mapas com máscaras	58
2.7	O significado das anisotropias	61
3	Detectabilidade da topologia cósmica	65
3.1	Introdução	65
3.2	Definindo a detectabilidade	67
3.3	Modelos Λ CDM: Condições para detectabilidade da topologia	73
3.3.1	Um modelo simples	73
3.3.2	Método da reta secante	77
3.3.3	Método da reta tangente	79
3.3.4	As curvas de contorno e suas aproximações	80
3.4	O Gás de Chaplygin generalizado	85
3.4.1	Um fluido único	85
3.4.2	Detectabilidade da topologia em cosmologias de GCG	87
3.5	Considerações gerais sobre a detectabilidade	91
4	O problema inverso em topologia cósmica	94
4.1	Introdução	94
4.2	Círculos no céu e supernovas em modelos Λ CDM	96
4.2.1	O método de círculos no céu	97
4.2.2	O dodecaedro de Poincaré como candidato	100
4.2.3	Combinando dados: supernovas tipo SN Ia e a topologia cósmica	101
4.3	O problema inverso da detectabilidade no gás de Chaplygin generalizado	105
4.3.1	Escolhendo as variáveis	105
4.3.2	Combinando dados: supernovas tipo SN Ia e a topologia cósmica II	107
5	A forma local do Universo	110
5.1	Introdução	110
5.2	A forma local em espaços de curvatura positiva	112
5.2.1	Caracterizando as holonomias translacionais	112
5.2.2	Obtendo as geodésicas fechadas mais curtas	116

5.3	A forma local do Universo no limite inflacionário	120
5.3.1	A inflação	120
5.3.2	A forma local do Universo no limite inflacionário	123
5.3.3	Consequências observacionais	129
6	Efeito do movimento próprio de nossa galáxia na detecção de círculos no céu	131
6.1	Introdução	131
6.2	A aberração local da radiação cósmica	132
6.3	Consequências para buscas futuras	138
	Considerações finais	141
A	O espectro de potência angular e a física das anisotropias	144
B	A classificação das 3-variedades	150
B.1	$\widetilde{M} = \mathbb{R}^3$ (curvatura $K = 0$)	150
B.2	$\widetilde{M} = \mathbb{S}^3$ ($K > 0$)	153
B.3	$\widetilde{M} = \mathbb{H}^3$ ($K < 0$)	155
C	Curvas de confiança e observações de supernovas	157
C.1	Observações e regiões de confiança	157
C.2	A distância-luminosidade	160
C.3	Supernovas SN Ia como velas-padrão	161
D	Teoria dos números aplicada	163
D.1	Alguns resultados da teoria de números	164
D.1.1	Operações em módulo, e congruências	164
D.1.2	O teorema de Euler, e o pequeno teorema de Fermat	164
D.1.3	Frações continuadas	165
D.2	Aplicação a espaços-lente	168
D.2.1	Grupos finitos de simetrias de $\mathbb{S}^3$ e os espaços-lente	168
D.2.2	A geodésica fechada mais curta, e a detectabilidade da topologia . .	171
D.2.3	O subgrupo das holonomias translacionais	173
	Bibliografia	175

Lista de Tabelas

3.1	Aproximações lineares das curvas de contorno para variedades hiperbólicas.	81
3.2	Cotas para o valor mínimo de Ω_0 que garante a indetectabilidade em variedades esféricas	83
3.3	Cotas superiores de Ω_0 para a indetectabilidade da topologia cósmica . . .	83
3.4	Condições necessárias para detectabilidade (ou suficientes para indetectabilidade) de variedades curvas em modelos Λ CDM	84
3.5	Condições suficientes de detectabilidade para variedades curvas em modelos GCG	91
5.1	Classificação dos espaços-lente $L(p, q)$, seus homeomorfos e subgrupos translacionais.	115
5.2	Condições suficientes de detectabilidade para espaços-lente em modelos Λ CDM	118

Lista de Figuras

1.1	Anisotropias de temperatura da radiação cósmica de fundo	18
1.2	Classes de homeomorfismos em superfícies compactas orientáveis	26
1.3	2-Toro em $\mathbb{S}^3$, e o seu domínio fundamental	27
1.4	Espaço de recobrimento do 2-toro	29
1.5	Espaço de recobrimento da garrafa de Klein	30
2.1	O histograma de diferença $\Upsilon(\alpha)$	49
2.2	O σ -mapa de céu inteiro obtido do mapa ILC	53
2.3	O espectro de potência angular do mapa ILC	55
2.4	Multipolos do mapa ILC	56
2.5	σ -mapas para LILC, COADDED, OT e TOH	58
2.6	O efeito da máscara KP2 no histograma da diferença Υ	60
2.7	O mapa LILC após a aplicação da máscara KP2, e o σ -mapa resultante. . .	61
2.8	Espectro de potência do mapa LILC-KP2	62
3.1	Detectabilidade de imagens múltiplas em um Universo em expansão	70
3.2	Profundidade máxima de um levantamento em um modelo Λ CDM	74
3.3	Regiões de confiança no plano paramétrico $\Omega_{m0} - \Omega_{\Lambda0}$	76
3.4	Representação esquemática de uma curva de contorno do tipo $\chi_{obs} = r_{inj}$.	77
3.5	Representação esquemática das curvas de contorno $\chi_{obs} = r_{inj}^M$ para $\alpha =$ $-1, 0$ e ∞	89
3.6	O valor de χ_{obs} para diversos valores de $\Omega_0 = \Omega_{Ch0} + \Omega_{b0}$ em Universos curvos	92
4.1	Imagens múltiplas da última superfície de espalhamento, gerando pares de círculos identificados por uma isometria	98
4.2	Par de círculos antipodais na última superfície de espalhamento, e os ele- mentos do triângulo retângulo esférico ρ , r_{inj} e χ_{lss}	102
4.3	Regiões de confiança no plano paramétrico $\Omega_{m0} - \Omega_{\Lambda0}$ para a combinação de dados de SNs Ia com um vínculo topológico	104
4.4	Curvas de contorno de $\chi_{obs}(\Omega_0, \alpha) = r_{inj}$	106

4.5	Superposição no plano paramétrico $A - \alpha$ da região de confiança a 95% para dados de SNs Ia e as curvas de contorno $\chi_{obs} = r_{inj}$	107
4.6	Superposição da região de confiança a 95% no plano paramétrico $\Omega_0 - A$ para dados de SNs Ia e as curvas de contorno $\chi_{obs} = r_{inj}$	109
5.1	Detectabilidade de espaços-lente $L(p, q)$ em termos dos parâmetros p e q .	119
5.2	Variação no comprimento de geodésicas fechadas geradas por uma holonomia	126
6.1	Variações máximas e mínimas dos raios e posições dos centros dos círculos no céu com a mudança de referencial	135
6.2	Variação na posição dos centros dos círculos com a mudança de referencial	136
6.3	Variação no raio dos círculos com a mudança de referencial	136
6.4	Deslocamento angular para círculos de diferentes raios e centros	138
6.5	Efeito de deslocamento de fase ao longo do perímetro dos círculos com a mudança de referencial	140
A.1	O espectro de potência da radiação cósmica de fundo obtido pelo WMAP .	145
A.2	A estrutura do espectro de potência da radiação cósmica de fundo	147
B.1	Domínios fundamentais das 3-variedades euclidianas compactas e orientáveis	152

Introdução

*And the end of all our exploring
Will be to arrive where we started
And know the place for the first time.*

T.S. Eliot

A máxima pela qual comumente nos conduzimos em nossos raciocínios é que os objetos dos quais não temos experiência assemelham-se àqueles dos quais a temos, que o que descobrimos ser mais usual é sempre mais provável, e que onde há oposição de argumentos devemos dar preferência aos que estão apoiados no maior número de observações passadas.

David Hume

Kepler, quando descobriu as leis do movimento elíptico dos planetas que hoje levam o seu nome, provavelmente já sabia da sua natureza revolucionária. A primeira vista, trocar órbitas circulares por órbitas elípticas pouco excêntricas poderia parecer um ligeiro aprimoramento nos detalhes de uma teoria bem-sucedida. Mas isto representava uma quebra fundamental do paradigma que afirmava ser o movimento dos corpos celestes perfeito (i.e., circular) e imutável, e intrinsecamente diferente dos imperfeitos e transientes movimentos na Terra. Mesmo após a sua contestação por Copérnico, o modelo Ptolomaico, com suas inúmeras categorias de movimentos circulares sobrepostos, era geralmente visto com entusiasmo, como a mais perfeita e precisa descrição do cosmos.

Os cálculos feitos por Kepler com base nas observações de Brahe certamente excluía um modelo geocêntrico com órbitas circulares, mas este modelo já havia sido descartado

na antiguidade. Muito mais importante foi a exclusão definitiva do modelo Ptolomaico. Por si só isto não seria suficiente para excluir completamente todo e qualquer modelo geocêntrico. Quaisquer que fossem as observações, seria sempre possível acrescentar mais epiciclos no modelo de Ptolomeu, até que este se encaixasse novamente nas observações, dentro da precisão observacional da época. De fato, foi exatamente desta forma que este modelo se desenvolvera.

Mas Kepler foi além. Seus cálculos permitiam excluir não só o modelo de Ptolomeu, mas também o modelo misto de Brahe (onde os planetas orbitam o Sol, que orbita a Terra), e até o modelo heliocêntrico com órbitas circulares de Copérnico. Depois de muita hesitação e dúvida, Kepler descobriu que um modelo heliocêntrico com órbitas elípticas se encaixava muito bem às observações.

A esta altura, como ele poderia determinar qual a melhor teoria? As observações eram compatíveis com dois modelos distintos: um modelo heliocêntrico com órbitas elípticas, e um modelo geocêntrico com epiciclos adicionais. Qual deles era o ‘melhor’? Se partíssemos da hipótese segundo a qual os movimentos celestes devem ser sempre composições de movimentos circulares, aceita quase axiomáticamente pelos estudiosos desde a época helenística, então teríamos que concluir que Kepler e Brahe, longe de provar o heliocentrismo, o refutaram, pelo menos na sua forma mais simples (lembrando que a simplicidade era um dos grandes atrativos da teoria de Copérnico).

Kepler, porém, ao se ater à hipótese heliocêntrica, encontrou mais do que simplesmente órbitas elípticas. Existiam duas outras leis, relacionando o tamanho, velocidade e período das órbitas, que se aplicavam ao movimento de todos os planetas. Os raios e períodos dos vários ciclos e epiciclos ptolomaicos eram essencialmente arbitrários, escolhidos a posteriori para se ajustarem às observações. O novo modelo de Kepler tinha a vantagem de unificar um grande conjunto de observações por meio de algumas poucas leis fenomenológicas simples e elegantes. Uma explicação mais fundamental para estas leis não tardou a vir. Inspirado diretamente por Kepler (e Galileu), Isaac Newton lançou os alicerces da física moderna com sua teoria da gravidade universal, que se aplicava igualmente ao céu e à Terra. As esferas de cristal e os epiciclos subitamente se tornaram tão anacrônicos como descrição do Universo quanto os proverbiais elefantes nas costas de uma tartaruga. Nada mal para a consequência de um ligeiro aprimoramento nos detalhes de uma teoria bem-sucedida.

A digressão acima ilustra um ponto recorrente na história da cosmologia. Em quase todas as épocas desde os primórdios das ciências naturais na Grécia antiga, os estudiosos dispunham de um modelo que explicava adequadamente quase todas as observações disponíveis. Acreditavam portanto dispor de uma descrição razoavelmente precisa e com-

pleta para o Universo. Pequenas discrepâncias não eram (ou são) ignoradas de todo, mas em vista do sucesso da teoria geralmente aceita, era razoável supor que desapareceriam após ligeiras correções. As grandes revoluções no entendimento do cosmo neste período tiveram origem em tentativas de explicar tais discrepâncias por meio da ruptura com um ou mais fundamentos do paradigma aceito (muitas vezes de forma implícita) até então. Foi assim quando Kepler prescindiu de órbitas circulares, ou quando Newton formulou explicitamente leis de aplicação universal acima e abaixo da 'esfera lunar', ou também quando Einstein modificou a extraordinariamente bem-sucedida dinâmica newtoniana com base em um novo conjunto de postulados, e ainda quando Hubble mostrou que o Universo não era estático.

Nas últimas décadas, a cosmologia baseada na hipótese de um Big Bang quente com sua dinâmica dada pela teoria da relatividade geral tem sido extremamente bem-sucedida em explicar precisamente uma grande quantidade de observações, que vão desde a abundância relativa de elementos leves até a estrutura em grandes escalas do Universo atual, passando pela natureza térmica e as anisotropias da radiação cósmica de fundo. O recente surgimento da chamada cosmologia de precisão só fez confirmar os fundamentos deste modelo, com uma acuidade sem precedentes. Mas ainda há alguns pontos desconhecidos, e ligeiras discrepâncias entre teoria e observações. Por exemplo, aparentemente existem pequenos desvios na gaussianidade e anisotropia estatística da radiação cósmica de fundo. E não sabemos exatamente do que consistem as chamadas matéria e energia escuras, ou a origem do campo escalar que supostamente gerou a inflação. Ou qual é exatamente a topologia do Universo. Talvez estes problemas tenham soluções relativamente simples dentro do paradigma vigente. Mas é também possível que sejam indícios de uma nova física, e/ou resultar em um entendimento qualitativamente mais profundo do cosmo.

Atualmente, existem fortes razões para acreditar que o Universo é, em grandes escalas, aproximadamente igual em todos os pontos, e em todas as direções, e bem descrito pela teoria da relatividade geral de Einstein. Neste contexto, a geometria das seções espaciais do espaço-tempo admite diversas topologias diferentes. É muito comum pressupor que esta topologia é sempre a mais simples possível (onde por exemplo toda curva fechada pode ser sempre contraída em um ponto). Mas não há nenhum imperativo observacional ou teórico para tal hipótese. Sabemos que a relatividade, apesar de todos os seus triunfos, também tem limitações. Em particular, como uma teoria geométrica local, ela não fixa como o espaço se conecta globalmente, o que é uma maneira coloquial de dizer que a relatividade determina localmente a geometria, mas não a topologia. Mais do que uma limitação, este fato indica que a topologia cósmica é, possivelmente, a janela para alguma teoria mais fundamental capaz de fixá-la. Em termos observacionais quantitativos, a pre-

sença de uma topologia não trivial para o Universo não seria particularmente importante. Mas a sua existência, e natureza exata, seriam as primeiras evidências diretas de uma física transrelativística. A questão que se impõe é portanto: o que sabemos, ou podemos descobrir por observações, sobre a topologia cósmica? É uma questão que deve levar em consideração uma característica do estudo do Universo como um todo.

A cosmologia é, assim como a história, uma área do conhecimento humano que tem a distinção de ter como objeto de estudos a evolução temporal um único elemento, por definição irreproduzível, dentro do qual os próprios investigadores estão inseridos. E, assim como a história, a cosmologia consiste em bem mais do que mera tabulação seqüencial de eventos importantes. Almejamos explicar por que o Universo é exatamente como o observamos, e não de outra forma, usando um conjunto de leis universais aplicáveis em todos os lugares e em todas as épocas. Mas nossa capacidade de observação é limitada por nossa posição no espaço e no tempo. Nada garante que podemos, ou jamais poderemos, ver todo o Universo. Assim, o conhecimento do cosmos, embora baseado em observações, deve ser sempre temperado pelo entendimento das limitações do que é possível a conhecer.

Uma questão fundamental ainda em aberto na cosmologia é saber se o Universo é finito ou infinito. De forma mais geral, podemos nos perguntar: qual é a forma do Universo? Como não existe, por definição, um observador externo a observá-lo, esta questão demanda uma certa sutileza para ser bem posta. Um bom começo é reformular ligeiramente a pergunta: de quantas maneiras diferentes dois pontos podem ser ligados? Em uma folha de papel estendida sobre uma mesa, dois pontos podem ser ligados por todo tipo de curvas complicadas; mas estas podem ser sempre deformadas suavemente no segmento de reta que representa a menor separação entre os pontos. Mas a mesma folha, quando enrolada sobre si mesma, forma um cilindro. Dois pontos em um cilindro podem ser ligados por curvas que dão zero, uma ou mais voltas em torno de seu perímetro; e para cada curva este número de voltas não pode ser alterado por qualquer deformação suave. Em linhas gerais, é este tipo de estrutura que determina o que chamamos de forma do Universo. A topologia cósmica consiste exatamente no estudo de tais propriedades de conectividade globais do espaço. Mas note que, no exemplo acima, para diferenciar entre uma folha infinita de papel e um cilindro, é preciso ser capaz de observar pelo menos uma volta completa deste último. Da mesma maneira, em princípio existe uma diferença entre a forma do Universo, e os seus elementos que somos capazes de deduzir a partir de observações, que chamamos de forma local. É importante portanto saber se, e sob quais condições, observações astrofísicas serão capazes de decidir esta questão.

O presente trabalho é uma contribuição ao entendimento de como e quando a topologia cósmica pode ser detectável, e das conseqüências observacionais de tal detecção. Ele se

insere em uma linha de pesquisa que pode ser resumida da seguinte forma:

- Quais discrepâncias existem entre os modelos cosmológicos vigentes e as observações? A topologia cósmica é uma boa explicação para elas? Para discrepâncias com causas não-topológicas, qual o seu impacto na detectabilidade da topologia?
- O que as observações podem nos dizer sobre a topologia cósmica? Quais topologias seriam detectáveis, sob que condições e com quais métodos?
- O que a topologia poderia nos dizer sobre o Universo? Se a topologia cósmica for detectada, em que medida seria possível restringir os demais parâmetros dos modelos cosmológicos?
- O que a topologia pode dizer sobre si mesma? Se a topologia for detectável, que forma terá? O que poderemos deduzir sobre a topologia como um todo a partir de seus efeitos detectáveis?
- Quais efeitos poderiam impedir ou dificultar a detecção da topologia, e o que fazer para contorná-los?

Tendo em vista estas questões, a tese está organizada em seis capítulos, que seguem o raciocínio descrito acima, e não a sequência cronológica dos trabalhos publicados. No Capítulo 1 são apresentados os tópicos fundamentais para o entendimento da topologia cósmica em geral e dos resultados mostrados nos capítulos subsequentes em particular. Procurei criar uma apresentação auto-contida que permita a um leitor familiarizado com princípios da cosmologia relativística entender o texto sem ter que consultar fontes externas; ao mesmo tempo, o capítulo é focado nos resultados úteis para o entendimento do presente trabalho, e não pretende ser um texto didático para uso geral.

No Capítulo 2 estudamos uma das possíveis discrepâncias entre teoria e observação na cosmologia. É apresentado um método que desenvolvemos para detectar de forma direcional e quantitativa desvios da anisotropia estatística na radiação cósmica de fundo. Mostramos que o céu de microondas possui uma direção preferencial de máxima anisotropia que é estatisticamente significativa.

No Capítulo 3 apresentamos de uma forma sistemática as condições de detectabilidade das topologias compatíveis com um Universo quase plano em termos dos parâmetros de modelos cosmológicos. Estudamos em particular modelos do tipo Λ CDM e gás de Chaplygin generalizado. No primeiro caso, desenvolvemos ainda um método analítico para obter condições suficientes para a detectabilidade da topologia.

Estudamos o problema inverso da detectabilidade da topologia cósmica no Capítulo 4, e obtemos restrições numéricas sobre parâmetros cosmológicos ao combinar as consequências observacionais de uma possível detecção da topologia cósmica com dados advindos da observação de supernovas do tipo SN Ia.

O Capítulo 5 é dedicado ao estudo da chamada forma local do Universo, e contém alguns dos resultados mais interessantes desta tese. Chegamos a conclusões bastante gerais sobre quais elementos da topologia serão detectáveis para um observador genérico, em particular no limite imposto sobre a curvatura do espaço por teorias inflacionárias. Estes resultados permitem simplificar significativamente eventuais buscas por evidências de uma topologia cósmica não trivial.

Finalmente, no Capítulo 6, tratamos de um efeito que pode potencialmente complicar a detecção da topologia cósmica: o movimento próprio de nossa galáxia em relação ao fluxo comóvel de Hubble. Caracterizamos de forma completa as consequências deste efeito sobre a detecção dos chamados círculos no céu, um dos principais métodos propostos para a detecção da topologia cósmica.

A tese é fechada com quatro apêndices. O Apêndice A discute brevemente a análise do espectro angular de potência da radiação cósmica de fundo, enquanto o Apêndice B trata da classificação das formas espaciais. O Apêndice C trata de forma sucinta da combinação de topologia com dados observacionais para restringir parâmetros de modelos cosmológicos. O Apêndice D consiste em uma breve introdução à teoria de números, e do seu uso para provar os resultados apresentados no Capítulo 5.

Este trabalho não tem a pretensão de resolver a questão da topologia cósmica de forma definitiva, ou de exaurir o tema. É uma área vasta e multifacetada, na qual muito já foi feito anteriormente, e muito ainda há por fazer. Espero porém que esta contribuição nos ajude a entender melhor o que podemos esperar detectar, e sob quais circunstâncias; e também quais seriam algumas das consequências de uma tal detecção.

Ao longo do texto, me refiro à autoria dos resultados obtidos sempre no plural. A preferência por 'nossos' resultados em detrimento de 'meus' resultados não é firula estilística, ou falsa modéstia. Todos os resultados aqui apresentados que já foram publicados são oriundos de colaborações envolvendo a mim, meu orientador, e um ou mais co-autores. Foram esforços coletivos pelos quais só posso assumir crédito parcial. E, de fato, este processo colaborativo foi possivelmente a parte mais gratificante do trabalho. Tive ao todo sete colaboradores, de seis intuições e quatro países de origem diferentes, uma rica diversidade de opiniões, formações e personalidades. Em ciência, como em muitas outras áreas, a experiência prática ganha trabalhando é muitas vezes mais útil que o aprendizado teórico; neste aspecto, posso afirmar que tive sete excelentes professores, sem os quais a

presente tese não seria possível. Obviamente, eventuais erros e omissões que porventura aqui existam são de minha inteira e única responsabilidade.

Capítulo 1

Preliminares

A matemática é a linguagem na qual Deus escreveu o Universo.

Galileu Galilei

Para aqueles que não conhecem matemática é difícil expressar a profunda beleza da natureza. Se você quer aprender sobre a natureza, é necessário aprender a linguagem na qual ela fala.

Richard P. Feynman

Os matemáticos são como os franceses: Sempre que alguém lhes diz alguma coisa, eles a traduzem para sua própria língua, e imediatamente ela se torna algo totalmente diferente.

Johann Wolfgang von Goethe

1.1 Introdução

As seções seguintes consistem de rápidas revisões de tópicos tais como cosmologia relativística ([47], [111], [128], [107]); radiação cósmica de fundo ([128] e [150]); topologia

([140], [75] e [105]), e em particular topologia em 3 dimensões ([169] é o texto padrão); e a procura pela topologia cósmica ([88], [93], [107], [62], [138]). Nosso objetivo é fixar conceitos fundamentais a serem usados mais adiante, e estabelecer o contexto físico e matemático relevante para esta tese. São revisões portanto seletivas, e de forma alguma completas ou matematicamente auto-suficientes. Existem vários trabalhos excelentes que tratam destes tópicos em qualquer nível de complexidade desejada. As referências citadas acima são uma seleção pessoal de fontes que considero bons pontos de partida para cada um dos tópicos mencionados.

1.2 Relatividade, e cosmologia relativística

1.2.1 Relatividade, beleza e universalidade

Possivelmente o exemplo conhecido mais perfeito da surpreendente conexão entre beleza e realidade física, a teoria especial da relatividade, proposta por Einstein em 1905, tem como ponto de partida duas premissas: O princípio de relatividade, segundo o qual as leis físicas são as mesmas em todos os referenciais inerciais; e a invariância da velocidade da luz no vácuo, c . Como consequência destes dois postulados, surge um arcabouço teórico que unifica geometricamente as noções de tempo e espaço, agora indissociáveis no conceito de espaço-tempo. E é nesta aparente inexorabilidade, juntamente com sua simplicidade essencial e universalidade de aplicação, que reside a beleza da Relatividade.

Seria natural (ou pelo menos hoje assim nos parece) procurar estender a relatividade especial para observadores em referenciais acelerados, sujeitos a campos gravitacionais, e produzir uma explicação geométrica natural para a gravidade. E foi de fato o que Einstein fez, dez anos depois, quando propôs a teoria geral da relatividade.

Assim como a relatividade especial, a relatividade geral também se baseia em algumas poucas premissas, fundamentadas em poderosas intuições físicas que até hoje aceitamos como naturais. Assim, admitimos que:

- As leis físicas são as mesmas para quaisquer referenciais (inerciais ou não): O princípio geral da relatividade.
- As leis físicas não devem depender do sistema de coordenadas usadas para expressá-las: O princípio de covariância.
- Localmente, a teoria especial da relatividade se aplica para todos os observadores: O princípio da invariância de Lorentz local.

Para obter uma teoria física aplicável, porém, é necessário lançar mão de hipóteses adicionais (essencialmente relacionadas com a maneira como gravidade e movimento são descritos matematicamente), que embora bastante plausíveis, não são axiomáticas. Assim, presumimos ainda que:

- A geometria do espaço-tempo é descrita por uma variedade pseudo-riemanianna (i.e., uma variedade diferenciável $\mathcal{M}$ na qual em cada ponto o espaço vetorial tangente $T_p(M)$ é dotada de um (pseudo)-produto interno.) Essencialmente, o espaço-tempo é curvo, e existem sempre bases locais em termos das quais a métrica é do tipo Minkowski.
- O movimento inercial (i.e., sujeito apenas à gravidade) de partículas-teste segue geodésicas tipo tempo (para partículas com massa) ou do tipo nulo (para partículas sem massa).
- O conteúdo material do espaço-tempo, descrito por seu tensor de matéria-energia $T_{\mu\nu}$, cria a curvatura do espaço-tempo segundo as equações de campo de Einstein

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \kappa T_{\mu\nu} - \Lambda g_{\mu\nu} , \quad (1.1)$$

onde $g_{\mu\nu}$ é o tensor métrico, Λ a chamada constante cosmológica, cujo valor é em princípio arbitrário, e $R_{\mu\nu}$ e R contrações sucessivas do tensor de curvatura de Riemann $R_{\alpha\mu\nu\beta}$. $\kappa = 8\pi G/c^4$. Daqui em diante escolheremos unidades de modo que $c = 1$.

A parte geométrica,

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} , \quad (1.2)$$

é chamada tensor de Einstein.

A conservação de energia-momento pode ser escrita de forma covariante como

$$T^{\mu\nu}{}_{;\nu} = 0, \quad (1.3)$$

e é assegurada pela equação (1.1), já que, usando as identidades de Bianchi, é possível mostrar que

$$G^{\mu\nu}{}_{;\nu} = 0 , \quad (1.4)$$

onde o símbolo “ $;$ ” denota a derivada covariante.

Uma descrição completa do espaço-tempo deve levar em conta a natureza de seu conteúdo material (determinado pelo tensor de energia-momento), e a dinâmica que rege o seu movimento. A equação de Einstein relaciona explicitamente estes dois elementos. Formalmente, consiste em um conjunto de equações diferenciais não-lineares, cujas soluções

correspondem ao tensor métrico do espaço-tempo. Como afirma o aforisma atribuído a Wheeler, a matéria diz ao espaço como se curvar, e o espaço diz à matéria como se mover.

Existem formulações equivalentes da relatividade, assim como teorias alternativas. Mas a teoria da relatividade tem se mostrado em excelente concordância com diversas observações que poderiam contestá-la (ou quase todos, aparentemente; a anomalia no movimento de sondas enviadas para além do sistema solar ainda carece de explicação, e talvez seja evidência de uma física pós-relativística [120]). Daqui em diante, vamos sempre supor que a teoria geral da relatividade permite uma descrição precisa das interações gravitacionais em escalas não-quânticas.

1.2.2 O Universo como objeto de estudo

A relatividade geral é uma teoria métrica local, ou seja, ela determina localmente a geometria do espaço-tempo a partir da distribuição de matéria-energia. Se por um lado isto é uma limitação (um exemplo bastante típico: ela não determina, embora possa restringir a topologia das seções espaciais), ela também permitiu pela primeira vez a descrição do Universo como um todo por meio da linguagem matemática. O objeto no caso é uma variedade pseudo-riemanianna $(\mathcal{M}, g)$ que tomamos como espacialmente homogênea e isotrópica em grandes escalas, e cuja métrica é descrita pelas equações de campo de Einstein (1.1).

As hipóteses de homogeneidade e isotropia nos permitem definir as propriedades de um fluido universal somente em uma pequena vizinhança, para só então generalizar nossas hipóteses para obter daí um tensor métrico aplicável a todo o espaço-tempo. Considere então um 'fluido universal', de composição inicialmente arbitrária. Podemos definir a 4-velocidade do centro de massa do conteúdo de uma vizinhança para cada ponto da variedade. Temos assim uma família de linhas de mundo, cujos componentes definem um sistema de coordenadas não-singular

$$u^\alpha = \frac{dx^\alpha}{d\tau} , \quad (1.5)$$

$$u_\alpha u^\alpha = -1 . \quad (1.6)$$

É possível então escrever a representação de um tensor energia-momento $T_{\alpha\beta}$ genérico em termos do campo de quadrivelocidades $\mathbf{u}$, e do projetor ortogonal a $\mathbf{u}$, $h_{\alpha\beta} = g_{\alpha\beta} + u_\alpha u_\beta$ (assim $h_{\alpha\beta} u^\alpha = 0$). Define-se ainda $\rho = (T_{\alpha\beta} u^\alpha u^\beta)$ como a densidade total de energia, $p = (1/3)(T_{\alpha\beta} h^{\alpha\beta})$ como a pressão isotrópica, $\pi_{\alpha\beta} = T_{\gamma\delta} (h^\gamma_\alpha h^\delta_\beta - h^\gamma_\beta h^\delta_\alpha)$ como a pressão anisotrópica (em termos newtonianos, a resposta na direção α à pressão exercida

na direção β), e $q^\alpha = -T_{\beta\gamma} u^\beta h^{\gamma\alpha}$ como o fluxo relativístico de energia . Note que as seguintes propriedades de simetria e ortogonalidade podem ser obtidas diretamente das definições

$$q_\alpha u^\alpha = 0, \quad \pi^\alpha{}_\alpha = 0, \quad \pi_{\alpha\beta} = -\pi_{\beta\alpha}, \quad \pi_{\alpha\beta} u^\beta = 0. \quad (1.7)$$

Uma excelente discussão sobre a relação entre hipóteses de simetria e modelos cosmológicos pode ser encontrada nas notas de curso de Ellis e Van Elst [49]. Combinando as quantidades definidas acima, é fácil verificar que $T_{\alpha\beta}$ pode ser decomposto em seus componentes ortogonais

$$T_{\alpha\beta} = \rho u_\alpha u_\beta + q_\alpha u_\beta + u_\alpha q_\beta + p h_{\alpha\beta} + \pi_{\alpha\beta}. \quad (1.8)$$

O fluido descrito acima é absolutamente genérico, mas podemos agora impor-lhe vínculos físicos. Em particular, para que a isotropia local se mantenha é preciso que a congruência de linhas de mundo definidas por $\mathbf{u}$ seja irrotacional ($q_\alpha = 0$), e não sofra deformações anisotrópicas ($\pi_{\alpha\beta} = 0$). Estas condições caracterizam um fluido perfeito,

$$T_{\alpha\beta} = \rho u_\alpha u_\beta + p h_{\alpha\beta}. \quad (1.9)$$

Sob condições bastante gerais, a evolução de um fluido relativístico pode ser descrita localmente pela equação de Raychaudhuri (veja [137]), mas no caso de um fluido perfeito localmente isotrópico e homogêneo é fácil ver que uma esfera infinitesimal só pode se contrair ou expandir uniformemente, sem rodar ou se distorcer.

Em geral, um vínculo físico extra é imposto a um fluido perfeito, na forma de uma equação de estado que relaciona pressão e densidade; $p = p(\rho, x_0)$. Casos importantes incluem a 'poeira', ou matéria escura (fria), com $p = 0$; a radiação eletromagnética, com $p = \frac{1}{3}\rho$; e a constante cosmológica ou energia escura, $p = -\rho$. De modo mais geral, são normalmente impostas algumas restrições às equações de estado possíveis (as condições de energia fraca, nula ou forte; respectivamente $\rho > 0$, $\rho + p > 0$ e $\rho + 3p > 0$; estabilidade da matéria $\partial p / \partial \rho \geq 0$ e causalidade $\partial p / \partial \rho < 1$ são as mais comuns).

Podemos sempre encontrar localmente um referencial no qual $u^\alpha = \delta_0^\alpha$. Se a condição de isotropia espacial for igualmente válida para todos os pontos da variedade, a congruência u define de forma consistente um sistema de coordenadas no qual a condição $u^\alpha = \delta_0^\alpha$ seja identicamente válida, de modo que sempre existirá localmente um referencial (dito comóvel) em repouso em relação a u , para o qual x^0 representa o tempo próprio. Isto equivale a dizer que existe uma densidade e uma pressão relativística definida identicamente em todo o espaço, com $T_{\alpha\beta} = \text{diag}\{\rho, p, p, p\}$. Neste caso é possível

mostrar que no referencial comóvel o espaço-tempo terá uma métrica do tipo Friedmann-Lemaître-Robertson-Walker (FLRW; muitas vezes alguns dos nomes são omitidos, em geral seguindo preferências nacionais)

$$ds^2 = -dt^2 + a^2(t) [d\chi^2 + f^2(\chi)(d\theta^2 + \sin^2 \theta d\phi^2)] . \quad (1.10)$$

O tempo t , o tempo próprio dos observadores comóveis, pode ser considerado um tempo cosmológico universal. A variedade $\mathcal{M}$ pode então ser naturalmente decomposta em suas partes temporal e espacial; $\mathcal{M} = \mathbb{R}^+ \times M$. As seções espaciais M , ortogonais ao fluxo de Hubble $u^\alpha = \delta_0^\alpha$, são variedades genuinamente riemaniannas (i.e., com um produto interno positivo-definido) com curvatura constante, e sua evolução é determinada por um fator de escala $a(t)$. Para curvatura nula, positiva e negativa ($k = 0, 1$ e -1), $f(\chi)$ é respectivamente χ , $\sin \chi$ e $\sinh \chi$. É comum supor que para cada um destes casos M seja simplesmente conexo, e corresponda (a menos de um fator de escala, ou homotetia), aos espaços euclidiano $\mathbb{R}^3$ (para $k = 0$), esférico $\mathbb{S}^3$ (para $k = 1$) e hiperbólico $\mathbb{H}^3$ (para $k = -1$). Mas a Relatividade Geral é uma teoria local, que restringe mas não determina a topologia de M . Como veremos, para cada um destes casos existe uma família de espaços quocientes $M = \widetilde{M}/\Gamma$ multiplamente conexos e potencialmente finitos (Γ é um grupo discreto e de ação livre de isometrias do espaço de recobrimento $\widetilde{M} = \mathbb{R}^3, \mathbb{S}^3$ ou $\mathbb{H}^3$). Nos dois últimos casos, $a(t)$ é o inverso do raio de curvatura de $\widetilde{M}$, i.e., $k/a(t)^2 = R/2$.

Usando a métrica FLRW e o tensor de energia-momento de um fluido perfeito, obtemos as equações de Friedmann, que podem ser usadas para determinar o sinal de k e a evolução do fator de escala $a(t)$:

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho}{3} + \frac{\Lambda}{3} - \frac{k}{a^2} , \quad (1.11a)$$

$$\frac{\ddot{a}}{a} = \frac{\Lambda}{3} - 4\pi G \left(\frac{\rho}{3} + p\right) . \quad (1.11b)$$

Podemos definir uma quantidade chamada de densidade crítica, $\rho_c = \frac{3H^2}{8\pi G}$, que usamos para escrever a primeira equação de Friedmann em termos de parâmetros de densidade $\Omega_{X_i} = \rho_{X_i}/\rho_c$. Assim, se $\Omega_T = \sum \Omega_{X_i}$,

$$\Omega_T + \Omega_\Lambda = 1 + \frac{k}{H^2 a^2} ; \quad (1.12)$$

onde $\Omega_\Lambda = \rho_\Lambda/\rho_c = \Lambda/3\rho_c$.

Fica claro portanto que o parâmetro de densidade total $\Omega = \Omega_T + \Omega_\Lambda$ determina o sinal de k ; pois se $\Omega = 1$, > 1 ou < 1 , então a equação acima só pode ser satisfeita para k respectivamente igual a 0, 1 ou -1 .

A evolução do fator de escala $a(t)$ depende da equação de estado do fluido universal. A equação de conservação (1.3) do tensor de energia-momento de um fluido perfeito se reduz a

$$\dot{\rho} = -3H(\rho + p) . \quad (1.13)$$

No caso em que $p = p_X = \omega\rho_X$, a equação acima pode ser resolvida explicitamente

$$\rho_X = \rho_{X0} \left(\frac{a}{a_0} \right)^{-3(1+\omega)} . \quad (1.14a)$$

Os subscritos "0" denotam valores atuais. O desvio para o vermelho cosmológico de um fóton é simplesmente a razão entre os fatores de escala quando houve sua emissão, e detecção: $1 + z(t) = a_0/a(t)$. Em termos de z a equação acima é reescrita

$$\rho_X = \rho_{X0} (1 + z)^{3(1+\omega)} . \quad (1.15a)$$

A segunda equação de Friedmann indica que componentes cuja pressão obedece $p > -\frac{\rho}{3}$ contribuem para desacelerar a expansão do Universo; nesta categoria se incluem obviamente a matéria escura e a radiação, ou de modo mais geral, qualquer componente com pressão positiva. Uma constante cosmológica positiva, por outro lado, tem uma contribuição repulsiva. É possível definir uma densidade associada à constante cosmológica (ou, mais propriamente, à energia escura), $\rho_\Lambda = \frac{\Lambda}{8\pi G}$, com uma pressão associada $p_\Lambda = -\frac{\Lambda}{8\pi G}$. Estas definições preservam a estrutura das equações de Friedmann, de modo que os termos adicionais associados a Λ desaparecem, substituídos por um componente material com equação de estado $p = -\rho$. Usando (1.13), fica claro que o parâmetro de densidade deste componente permanece constante para qualquer valor de a : $\Omega_\Lambda = \text{const.}$ Para a matéria fria (escura ou bariônica), temos que $\omega_m = 0$, e para a radiação eletromagnética, $\omega_r = 1/3$. Portanto, $\Omega_m = \Omega_{m0} \left(\frac{a}{a_0} \right)^{-3}$ e $\Omega_r = \Omega_{r0} \left(\frac{a}{a_0} \right)^{-4}$. Torna-se claro então que se matéria fria, radiação e energia escura/constante cosmológica são os principais constituintes materiais do Universo, a contribuição da radiação tende a ser mais significativa no passado, e a da energia escura se torna dominante no futuro.

1.3 A radiação cósmica de fundo

1.3.1 A pré-história da RCF

Embora indícios claros da expansão do Universo existam desde a década de 1920, foi somente em 1948 que George Gamow e Ralph Alpher [5] sugeriram a existência de uma

radiação fóssil de microondas, térmica e isotrópica, oriunda de uma fase primordial na qual o Universo seria mais denso e quente do que hoje¹. Os resultados de Gamow e Alpher não eram bem conhecidos, até serem redescobertos independentemente por Robert Dicke e Yakov Zel'dovich, na década de 60. A existência teórica da radiação de fundo ganhou então maior aceitação, e em 1964 cientistas na Universidade de Princeton, liderados por Peter Roll e David Todd Wilkinson², começaram a construir um radio-observatório pioneiro inspirado pelas idéias de Dicke, com o objetivo de detectar tal radiação.

Em 1965, Arno Penzias e Robert Woodrow Wilson já haviam construído um instrumento semelhante, sem saber dos esforços paralelos em Princeton e alheios aos resultados teóricos que previam uma radiação de fundo em microondas. Com extremo cuidado e paciência, eles calibraram o seu aparato em várias frequências, apontando-o para regiões 'vazias' do céu. Após eliminarem todos os potenciais fatores de erros sistemáticos e contaminação (incluindo um casal de pombos que se alojara na antena), se depararam com um inexplicável ruído térmico, de $3,5\text{ K}$, que se recusava a desaparecer [131]. Finalmente, quase em desespero, Penzias e Wilson entraram em contato com o grupo de Princeton, que reconheceu [46] o ruído inconveniente como a radiação prevista havia quase 20 anos por Gamow e Alpher. Penzias e Wilson ganharam o prêmio Nobel por sua descoberta acidental, em 1968.

Durante o final da década de 1960, e início da década de 1970, a origem cosmológica do fundo de microondas foi repetidamente posta em dúvida (em particular por defensores de cosmologias de estado estacionário); mas com a confirmação do espectro absolutamente térmico da radiação, e o acúmulo crescente de evidências a favor de um modelo de "Big Bang" quente, a hipótese da origem cosmológica ganhou aceitação quase universal.

Até a década de 1990, porém, a radiação cósmica de fundo (doravante RCF) não era de grande valia para os cosmólogos, para além da confirmação da sua própria existência e natureza térmica. Acreditava-se que pequenas flutuações de densidade no Universo primordial, que seriam a origem de sua estrutura em grande escala que observamos hoje, deveriam originar anisotropias na temperatura da RCF; mas tais anisotropias, se existissem, não eram detectáveis pelos instrumentos disponíveis na época.

Em 1992, o observatório orbital COBE (COsmic Background Explorer) detectou e mapeou pela primeira vez as anisotropias previstas pela teoria, no que foi descrito por

¹Eles acreditavam ser esta uma explicação para a nucleogênese de todos os elementos mais pesados que o hidrogênio. A teoria universalmente aceita atualmente, da nucleossíntese estelar, foi desenvolvida por Fred Hoyle e colaboradores entre 1946 e 1954 [27].

²Por ocasião da morte de Wilkinson, o satélite MAP, (Microwave Anisotropy Probe, sobre o qual voltaremos a falar em breve), foi rebatizado WMAP (*Wilkinson* Microwave Anisotropy Probe) em sua homenagem.

Smoot como o nascimento da cosmologia de precisão [153]. Desde então, diversos outros instrumentos produziram mapas, parciais ou de céu inteiro, da RCF, culminando no satélite WMAP, do qual falaremos mais no capítulo seguinte. Para um relato mais completo da pré-história do estudo da RCF, veja [127], e também o sempre recomendável livro de Peebles, [128].

Em suas linhas gerais, o quadro que emerge deste conjunto de observações é essencialmente incontroverso. A RCF foi emitida quando o Universo se tornou frio o suficiente (cerca de 3000 K) para que a matéria bariônica se desionizasse e desacoplasse da radiação. Isto se deu aproximadamente 400.000 anos após o Big Bang, correspondente a um desvio para o vermelho cosmológico de $z = 1089$ (o que equivale a dizer que uma dada região do espaço se expandiu por um fator de 1089 desde então). A temperatura da radiação observada é inversamente proporcional ao fator de escala, e assim observamos hoje a RCF com uma temperatura da ordem de 3 K . As anisotropias observadas são a primeira vista estatisticamente isotrópicas, gaussianas e derivadas de um espectro de flutuações primordial invariante por escala. Existem pequenas discrepâncias em relação a todos estes itens (no próximo capítulo discutimos os desvios da isotropia estatística), mas o arcabouço fundamental permanece sólido. Medidas bastante diretas das posições dos picos acústicos também restringem significativamente uma série de parâmetros cosmológicos, notadamente a densidade total, Ω_0 , que é muito próxima de 1, o que corresponde a um Universo plano ou quase plano (investigamos as conseqüências de um Universo quase plano para a detecção de uma eventual topologia cósmica no Capítulo 5).

1.3.2 A análise das anisotropias

Como é usual, o ponto de partida para a análise das flutuações de temperatura, ou anisotropias, da RCF na esfera celeste é a sua decomposição em harmônicos esféricos, $Y_{\ell m}(\theta, \phi)$,

$$T(\theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\theta, \phi) , \quad (1.16)$$

onde a direção $\theta = 0$ é convencionalmente o pólo norte galáctico.

Toda a informação presente nos mapas de temperatura da RCF está contida nos coeficientes reais $a_{\ell m}$. E logo se torna claro que os valores destes coeficientes correspondem a três ordens de grandeza: aquele do dito monopolo, a_{00} , os componentes do dipolo a_{1-1} , a_{10} e a_{11} , e os demais multipolos.

Foi o monopolo que Penzias e Wilson efetivamente detectaram. É de longe o com-

ponente mais significativo³, e medições multiespectrais mostram corresponder a um dos mais perfeitos espectros de corpo negro já detectados na natureza, com uma temperatura de $2,725 \pm 0,001 \text{ K}$ [106].

Os componentes de dipolo tem um máximo na direção⁴ $l = 263,86^\circ$ e $b = 48,24^\circ$, e uma amplitude de 3.358 mK [42], cerca de 10^{-3} vezes menos intensa que a do monopolo. Acredita-se que praticamente todo o dipolo seja consequência do movimento próprio do observatório (que é a composição do movimento do satélite ao redor do Sol e da Terra, do Sistema Solar na Via Láctea, e desta em relação às demais galáxias, sendo este último o componente mais significativo), em relação ao referencial comóvel local, no qual se supõe que a RCF é térmica e aproximadamente isotrópica. O efeito Doppler resultante deste movimento induz um componente de dipolo em primeira ordem em $\beta = v/c$ (onde v é a velocidade do movimento próprio do observador), que é muito mais significativo do que o dipolo intrínseco da RCF (por um fator de 10^2), e indica que estamos nos movimentando a cerca de $1,23 \times 10^{-3}c$ na direção do máximo do dipolo.

O movimento próprio do observador também induz distorções menos significativas em multipolos de ordem superior. Analisamos as consequências destes efeitos para a detecção da topologia cósmica no Capítulo 6.

Após a remoção do monopolo (por definição homogêneo) e do dipolo, as anisotropias restantes, correspondentes a $\ell \geq 2$ (multipolos de ordem ℓ correspondem aproximadamente a escalas angulares da ordem de π/ℓ) têm amplitudes da ordem de 10^{-5} vezes a amplitude do monopolo (veja a Figura 1.1 para uma ilustração). Acredita-se que sejam consequência de flutuações adiabáticas de densidade no Universo primordial. Parece-nos justo afirmar que a análise destes componentes foi um grande triunfo para a cosmologia teórica, que teve suas premissas e alguns de seus modelos confirmados com considerável precisão. Um estudo detalhado destes resultados está fora do escopo da presente tese, mas apresentamos aqui um apanhado dos pontos mais relevantes para o que vem a seguir, a começar pelos aspectos mais gerais.

³Estritamente falando os instrumentos que efetivamente detectaram os demais componentes são insensíveis ao valor de a_{00} , pois são detectores diferenciais que medem a variação da radiação incidente vinda de duas direções diferentes. As medições mais precisas do monopolo disponíveis atualmente foram feitas pelo instrumento FIRAS (Far-InfraRed Absolute Spectrophotometer)[106], parte do satélite COBE.

⁴Usamos aqui coordenadas galácticas com o pólo norte galáctico em $b = 0^\circ$, o equador definido por $b = 90^\circ$, o centro galáctico em $l = 0^\circ$ e $l \in [0^\circ, 360^\circ)$ crescendo do oeste para o leste.

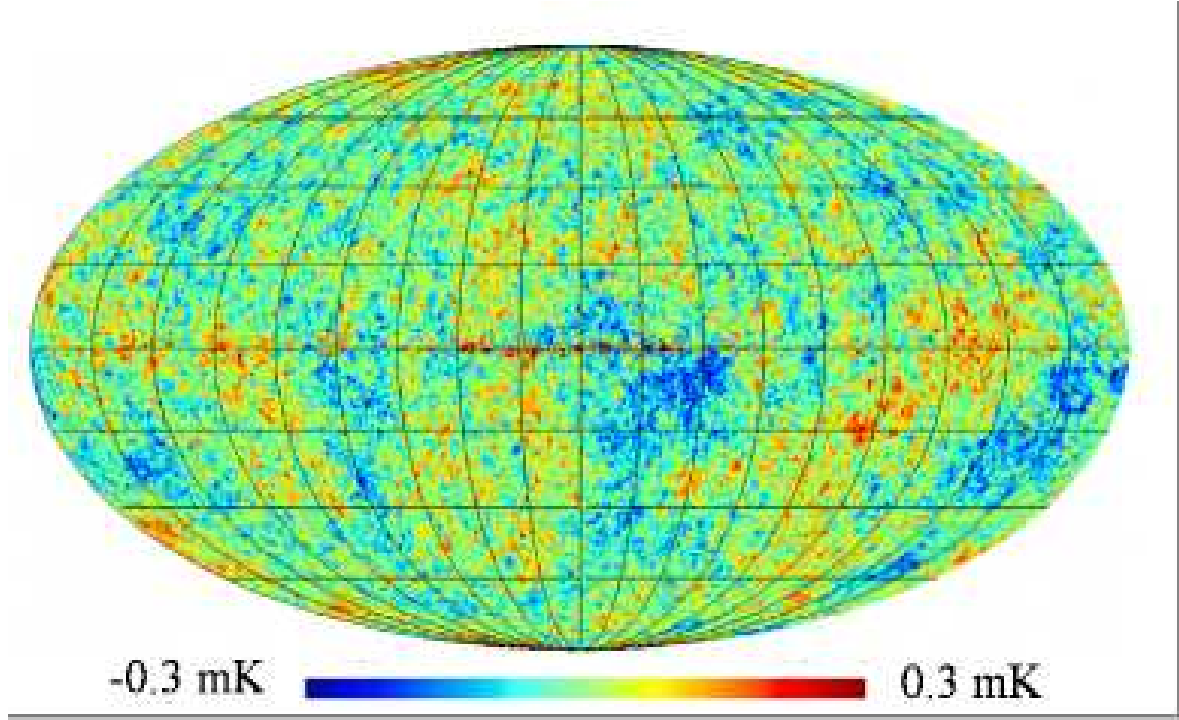


Figura 1.1: Anisotropias de temperatura da radiação cósmica de fundo, medidas pelo satélite WMAP. O componente de dipolo e a contaminação galáctica foram removidos.

1.3.3 Gaussianidade e isotropia estatística

Os modelos usuais inflacionários, nos quais um campo escalar decai lentamente ao longo de um poço de potencial suave (ditos 'slow roll'), implicam que as anisotropias da RCF devem ser gaussianas com alto grau de precisão (veja, e.g., [67] e [36]), ou seja, seu funcional de densidade de probabilidade [122] é dado por

$$P[\Delta T(\hat{n})] = \frac{1}{\mathcal{N}} \exp \left[-\frac{1}{2} \int \mathcal{C}^{-1}(\hat{n}, \hat{n}') \Delta T(\hat{n}) \Delta T(\hat{n}') d\Omega d\Omega' \right], \quad (1.17)$$

onde $\mathcal{N}$ é o fator de normalização e a covariância $\mathcal{C}(\hat{n}, \hat{n}')$ (a função de correlação de dois pontos) é

$$\mathcal{C}(\hat{n}, \hat{n}') = \langle \Delta T(\hat{n}) \Delta T(\hat{n}') \rangle. \quad (1.18)$$

A gaussianidade implica que $P[\Delta T(\hat{n})]$ é totalmente determinada pela covariância $\mathcal{C}(\hat{n}, \hat{n}')$, a partir da qual todas as demais funções de correlação de ordem superior podem ser obtidas.

Os harmônicos esféricos formam uma base ortonormal para funções suaves em $\mathbb{S}^2$, e portanto a gaussianidade definida acima implica em que os $a_{\ell m}$ são variáveis aleatórias gaussianas. De forma mais explícita,

$$\mathcal{C}^{-1}(\hat{n}, \hat{n}') \Delta T(\hat{n}) \Delta T(\hat{n}') = \sum_{\ell m; rs} a_{\ell m} a_{rs} Y_{\ell m}(\hat{n}) \mathcal{C}^{-1}(\hat{n}, \hat{n}') Y_{rs}^*(\hat{n}'), \quad (1.19)$$

e portanto

$$P(a_{\ell m}) = \frac{1}{\mathcal{N}} \exp \left[-\frac{1}{2} \sum_{\ell m; rs} a_{\ell m} a_{rs} \tilde{\mathcal{C}}_{\ell m; rs}^{-1} \right], \quad (1.20)$$

em que $\tilde{\mathcal{C}}_{\ell m; rs}^{-1} = \int Y_{\ell m}(\hat{n}) \mathcal{C}^{-1}(\hat{n}, \hat{n}') Y_{rs}^*(\hat{n}') d\Omega d\Omega'$ é um operador linear auto-adjunto (como $\mathcal{C}^{-1}(\hat{n}, \hat{n}')$ é auto-adjunto, $\tilde{\mathcal{C}}_{rs; \ell m}^{\dagger -1} = \tilde{\mathcal{C}}_{rs; \ell m}^{*-1}$). Novamente, a gaussianidade implica que a densidade de probabilidade $P(a_{\ell m})$ é totalmente determinada pela covariância $\tilde{\mathcal{C}}_{\ell m; rs} = \langle a_{\ell m} a_{rs} \rangle$.

Também é razoável supor, por simplicidade, que não existam eixos preferenciais, ou mais formalmente, que as funções de correlação das flutuações de temperatura são invariantes por rotações. Esta hipótese é chamada de isotropia estatística. É possível provar que em um mapa estatisticamente isotrópico de flutuações gaussianas em $\mathbb{S}^2$, toda informação sobre a RCF pode ser obtida a partir da covariância, e que esta é totalmente determinada pelo espectro de potência angular C_ℓ

$$C_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2, \quad (1.21)$$

com

$$\tilde{\mathcal{C}}_{\ell m; rs} = \delta_{\ell r} \delta_{ms} \langle C_\ell \rangle, \quad (1.22)$$

de onde é possível deduzir que

$$\mathcal{C}(\hat{n}, \hat{n}') = \frac{1}{4\pi} \sum_{\ell=1}^{\infty} (2\ell + 1) C_\ell P_\ell(\hat{n} \cdot \hat{n}'). \quad (1.23)$$

Tanto a gaussianidade quanto a isotropia estatística podem ser diretamente testadas em mapas da RCF (c.f. [52]), mas como dispomos de somente um Universo, só seria possível descartá-las na prática se os desvios fossem consideráveis (veja a discussão sobre variância cósmica, abaixo). Outros testes mais sofisticados indicam a possibilidade de desvios tanto da gaussianidade quanto da isotropia estatística (veja o Capítulo 2 para uma discussão do conceito de isotropia estatística, e de nossos resultados a respeito de possíveis desvios).

1.3.4 A variância cósmica

Como vimos acima, propriedades tais como gaussianidade e isotropia estatística são definidas em termos de uma função (ou funcional) de distribuição de probabilidades. Mas obviamente só na prática temos acesso a um céu de microondas, e portanto somente a uma realização desta distribuição.

Vimos ainda que se presumirmos gaussianidade e isotropia na RCF, os componentes $a_{\ell m}$ serão variáveis gaussianas independentes (de média zero), e as funções de correlação serão totalmente determinadas pelo espectro de potência angular. Logo, é útil estimar, a partir dos valores esperados para C_ℓ , qual a variância dos valores que seriam efetivamente medidos (veja e.g. [128] para maiores detalhes).

Por hipótese, cada um dos C_ℓ observados é uma variável aleatória resultante da soma de $2\ell + 1$ variáveis aleatórias gaussianas independentes. Sejam $\langle C_\ell \rangle$ os valores esperados do espectro de potência. Então C_ℓ segue a distribuição dada por

$$C_\ell \sim \frac{\sigma_{a_{\ell m}}^2}{2\ell + 1} \chi_k^2 = \frac{\langle C_\ell \rangle}{2\ell + 1} \chi_k^2 \quad (1.24)$$

onde χ_k^2 é a função qui-quadrado com $k = 2\ell + 1$ graus de liberdade para variáveis gaussianas X_m . Neste caso específico, $X_m = a_{\ell m}$, e $k = 2\ell + 1$

$$\chi_k^2 = \sum_{m=1}^k \frac{|X_m|^2}{\sigma_{X_m}^2} = \frac{1}{\langle C_\ell \rangle} \sum_{m=1}^{2\ell+1} |a_{\ell m}|^2 . \quad (1.25)$$

A densidade de probabilidade de se medir C_ℓ com o valor q é dada por

$$\rho(q; k) = \frac{1}{2^{k/2} \Gamma(k/2)} q^{k/2-1} e^{-q/2} . \quad (1.26)$$

É fácil calcular a média e variância desta distribuição:

$$\langle q \rangle_k = k , \quad (1.27a)$$

$$\sigma_{q;k}^2 = 2k . \quad (1.27b)$$

Escrevendo a variância de C_ℓ combinando (1.24) com (1.27a), temos

$$\begin{aligned} \sigma_{C_\ell}^2 &= \langle C_\ell^2 \rangle - \langle C_\ell \rangle^2 , \\ \sigma_{C_\ell}^2 &= \frac{\langle C_\ell \rangle^2 (\sigma_{q;2\ell+1}^2 + \langle q \rangle_{2\ell+1}^2)}{(2\ell + 1)^2} - \langle C_\ell \rangle^2 , \end{aligned} \quad (1.28)$$

$$\sigma_{C_\ell}^2 = \frac{2 \langle C_\ell \rangle^2}{2\ell + 1} . \quad (1.29)$$

O valor de $\sigma_{C_\ell}^2$ essencialmente indica o quão estatisticamente significativo é um desvio dos C_ℓ em relação à hipótese nula $\langle C_\ell \rangle$. Obviamente, quanto maior o valor de ℓ , maior será o número de componentes $a_{\ell m}$ e menor será a variância relativa. No capítulo seguinte usaremos uma variação deste cálculo da variância cósmica para calcular a significância estatística das anisotropias em grandes escalas angulares encontradas.

No Apêndice A, discutimos brevemente a física das anisotropias da RCF nas várias escalas angulares. Em particular, procuramos indicar de forma sucinta como o espectro angular de potência pode ser usado para restringir parâmetros dos modelos cosmológicos.

1.4 Topologia, geometria, e topologia cósmica

A topologia, além de ser uma legítima área da pesquisa matemática de interesse intrínseco, é também uma das grandes idéias unificadoras da matemática pura, provendo conexões entre áreas aparentemente díspares, tais como a álgebra, a teoria dos números e a teoria espectral. Conceitos topológicos surgem portanto naturalmente na matemática, e têm ampla aplicação na física.

Tais objetos e conceitos topológicos estão entre os mais abstratos e gerais que somos capazes de definir. Não é por outra razão que a topologia, mais até do que outras áreas da matemática pura, se expresse em termos de um conjunto intimidador de definições precisas e propriedades abstratas, que muitas vezes desafiam a capacidade humana de conceituação intuitiva. Em uma área tão ampla, vale a pena saber onde se quer chegar antes de começar a andar.

O presente trabalho não é um tratado matemático; a topologia aqui é um instrumento usado para investigar uma questão concreta a respeito do mundo físico: Qual a forma do Universo? Estudamos em particular a possibilidade do Universo possuir seções espaciais dotadas de uma topologia não trivial. Espaços topológicos genéricos prescindem de uma noção de distância (a base de uma visão geométrica do espaço), mas nossos objetos de estudo são variedades diferenciáveis tridimensionais, orientáveis e sem fronteiras, e dotadas de métricas positivo-definidas com curvatura constante. Dentre todos os espaços topológicos que podem ser definidos, esta é uma classe bem restrita. Tais variedades são também conhecidas como **formas espaciais**.

Não há, portanto, necessidade de se estudar a topologia em toda a sua generalidade. Tomaremos da teoria topológica geral somente aqueles conceitos necessários para caracterizar as variedades topológicas potenciais para as seções espaciais do Universo, no nível de generalidade que for preciso, e apresentaremos este ferramental matemático com o objetivo expresso de explorar suas conseqüências observacionais.

1.4.1 Fundamentos dos fundamentos da topologia

Um espaço topológico é um conjunto dotado de uma estrutura dada simplesmente pela definição da família de subconjuntos ditos 'abertos'. Mais precisamente,

Definição 1.1 Um **espaço topológico** é um conjunto S que contém uma família de subconjuntos $\mathcal{F}$ chamados de **conjuntos abertos**, dotados das seguintes propriedades:

i) $\mathcal{F}$ é fechado por interseções finitas, i.e., se A e $B \in \mathcal{F}$ então $A \cap B \in \mathcal{F}$.

ii) $\mathcal{F}$ é fechado por uniões arbitrárias i.e., se $\{A\} \subset \mathcal{F}$ então $\cup \{A\} \in \mathcal{F}$.

iii) $\emptyset \in \mathcal{F}$ e $S \in \mathcal{F}$.

O conjunto $\mathcal{F}$ é chamado de **topologia** de S . A um espaço topológico S dotado de topologia $\mathcal{F}$ denotamos $\langle S, \mathcal{F} \rangle$. Uma família de abertos $\mathcal{C} = \{C\} \subset \mathcal{F}$ é dita **cobertura** de S se $S \subset \cup \{C\}$.

As definições acima, embora abstratas, deixam claro que espaços topológicos são dotados de um conceito natural de conexão. A conexão entre dois pontos é determinada pelo subconjunto dos abertos que contém ambos. Mas este é um conceito elástico que nem sempre implica em uma definição unívoca de proximidade. Se um conjunto aberto que contenha um dado ponto $x \in S$ (o que chamamos **vizinhança** de x) também contém o ponto y , mas não o ponto z , então intuitivamente x poderia parecer estar mais 'próximo' de y do que de z . Mas nada garante que não exista uma segunda vizinhança de x que contenha z mas não y .

A noção intuitiva de que os abertos de uma topologia determinam como os pontos do espaço se conectam sugere por sua vez que seria útil definir aquelas transformações entre espaços topológicos que preservam tais conexões. Tais transformações, denominadas homeomorfismos, são por assim dizer os isomorfismos (i.e., as transformações que preservam as estruturas) dos espaços topológicos. Um objeto ou propriedade é dito topológico portanto quando é invariante por homeomorfismos:

Definição 1.2 Uma função $f : S \rightarrow \tilde{S}$ é dita **bijetiva** se $f(x) = f(y)$ implica em $x = y$ (injetividade), e se $\forall y \in \tilde{S} \exists x \in S \mid f(x) = y$ (sobrejetividade). Sejam $\langle S, \mathcal{F} \rangle$ e $\langle \tilde{S}, \tilde{\mathcal{F}} \rangle$ espaços topológicos. f é dita **contínua** se $\forall \tilde{A} \in \tilde{\mathcal{F}} \quad f^{-1}(\tilde{A}) \in \mathcal{F}$, onde f^{-1} denota a imagem inversa. Se f é contínua e se $\forall A \in \mathcal{F} \quad f(A) \in \tilde{\mathcal{F}}$, então f é dita **bicontínua**.

Uma bijeção bicontínua é denominada **homeomorfismo**.

Para determinar se uma transformação é homeomorfa, é necessário verificar se ela preserva a estrutura topológica definida pelo conjunto de todos os abertos. Ocasionalmente

(e, em particular, na Seção 1.4.3) se torna útil lançar mão de uma propriedade menos restritiva, que preserve somente a estrutura topológica em uma vizinhança de cada ponto. Assim,

Definição 1.3 *Um mapa contínuo $f : M \rightarrow W$ entre espaços topológicos M e W é um **homeomorfismo local** se, para cada ponto $x \in M$ existe uma vizinhança $V \subset M$ tal que $f(V)$ é um aberto de W e $f|_V : V \rightarrow f(V)$ é um homeomorfismo entre os espaços topológicos V e $f(V)$. Um homeomorfismo local é **sobrejetivo** se é bem definido em uma vizinhança U tal que $f : U \rightarrow W$.*

Claramente, a continuidade ou não de uma transformação depende da definição do conjunto dos abertos. A partir desta escolha um mesmo espaço pode ser dotado de diferentes topologias. Uma escolha mais restritiva dos abertos induz uma topologia mais 'fraca', na qual menos funções são contínuas. A inclusão de mais conjuntos abertos por outro lado torna a topologia mais 'forte', e aumenta a quantidade de funções contínuas. Uma maneira mais direta de classificar topologias é por sua capacidade de separar pontos. As únicas topologias relevantes para o presente trabalho são aquelas nas quais pontos distintos possuem vizinhanças distintas, i.e.,

Definição 1.4 *Um espaço topológico é dito **Hausdorff** (T_2) se para x e y , $x \neq y$, existem abertos A e A' tais que $x \in A$, $y \in A'$, e $A \cap A' = \emptyset$.*

Finalmente, assim como no caso dos espaços métricos, é útil definir uma subfamília, ou base, dos abertos a partir do qual todos os demais possam ser expressos.

Definição 1.5 *Em um espaço topológico $\langle S, \mathcal{F} \rangle$ denomina-se **base** de $\mathcal{F}$ uma família $\mathcal{B} \subset \mathcal{F}$ tal que $\forall A \in \mathcal{F}$ é possível escrever $A = \cup \{B\}$ para algum subconjunto $\{B\}$ de $\mathcal{B}$.*

Existe uma enorme liberdade para definir uma topologia, e textos introdutórios em geral precisam ser cuidadosos para tratar do tema em toda a sua generalidade (abordagens modernas e acessíveis podem ser encontradas em [75] e [105], com ênfase no moderno para o primeiro e no acessível para o segundo). Mas como dissemos acima, estamos interessados somente nas topologias potenciais das seções espaciais de variedades pseudo-Riemaniannas dotadas de uma métrica de FLRW. Localmente, estas seções (que são variedades riemaniannas) são localmente difeomorfas a $\mathbb{R}^3$. Por extensão, precisamos estudar somente os espaços topológicos localmente homeomorfos a $\mathbb{R}^3$.

A existência de homeomorfismos locais com $\mathbb{R}^n$ por si só não garante um espaço topológico bem comportado (i.e., em nosso contexto, potencialmente físico: passível de ser

dotado de métrica e de uma estrutura diferencial). Em um espaço métrico tal como $\mathbb{R}^n$, a definição usual de conjuntos abertos (A é aberto $\leftrightarrow \forall x \in A \exists \rho \in \mathbb{R}^+ \text{ t.q. se } d(x, y) < \rho \implies y \in A$) define a chamada topologia induzida. Tal topologia é trivialmente Hausdorff, e é fácil construir uma base enumerável, consistindo e.g. das bolas de raio $\epsilon_n = 10^{-n}$, $n \in \mathbb{Z}$) em torno de cada ponto $x \in \mathbb{Q}$ (a existência de uma base contável caracteriza as topologias ditas **segundo-contáveis**). Estas propriedades englobam a essência da estrutura dos abertos em um espaço métrico, e devem portanto ser preservadas em nossos espaços topológicos. Assim, os nossos objetos naturais de estudo são as chamadas variedades topológicas:

Definição 1.6 *Uma **variedade topológica** é um espaço topológico $\langle M, \mathcal{F} \rangle$, Hausdorff e dotado de base enumerável, com uma cobertura $\mathcal{C}$ tal que para qualquer $C \in \mathcal{C}$ existe um homeomorfismo $\varphi_C : C \rightarrow V_C \subset \mathbb{R}^n$.*

A maior parte das variedades topológicas que estudaremos a seguir (exceto pelas simplesmente conexas) serão compactas, o que corresponde à noção intuitiva de um objeto que pode existir em uma região finita do espaço.

Definição 1.7 *Uma variedade topológica é **compacta** se toda cobertura $\mathcal{C}$ possui uma subcobertura finita $\mathcal{C}' \subset \mathcal{C}$.*

Variedades compactas são sempre segundo-contáveis. De forma mais geral, as variedades segundo-contáveis são exatamente aquelas que podem ser imersas em $\mathbb{R}^n$, para algum n . A propriedade de Hausdorff, por outro lado, garante que limites de seqüências, se existirem, serão únicos, o que é fundamental para os processos de limite usados na definição de variedades diferenciáveis. Finalmente, tornando ligeiramente mais rigorosa a exigência de Hausdorff, é possível mostrar o chamado teorema de Urysohn-Tychonoff, que garante que tais variedades são metrizáveis.

A dimensão do espaço é uma propriedade topológica, então estudamos em particular as variedades topológicas em três dimensões, doravante 3-variedades. Note incidentalmente que caso tratássemos da topologia do *espaço-tempo*, estudaríamos variedades localmente homeomorfas ao espaço de Minkowski, e seria preciso relaxar a condição de Hausdorff, devido à existência de intervalos tipo luz.

1.4.2 Uma história de 2-toros

Em topologia, 2-toros são os melhores exemplos, e os piores exemplos. São casos simples, mas podem induzir generalizações simplistas. Com este aviso em mente, vamos usar o

humilde 2–toro para explicar algumas das noções básicas da topologia. No que se refere à topologia cósmica, as relações fundamentais entre topologia, geometria e observações são serão também ilustradas intuitivamente nesta seção.

Quando pode ser visualizada, a ação de um homeomorfismo é uma deformação suave de um objeto, que tem sua forma alterada sem rasgos ou emendas. Para uma dada família de espaços topológicos, os homeomorfismos definem classes de equivalência, que separam estes espaços em classes distintas de espaços topologicamente equivalentes. É natural portanto procurar classificar tais categorias da forma mais completa possível. Infelizmente, isto não pode ser feito de maneira geral, mas somente em alguns casos bem particulares. As superfícies compactas e orientáveis, por exemplo, são completamente categorizadas pelo número de buracos ("genus", plural "genii") que possuem. Assim, a esfera (homeomorfa a variedades sem alça, ou simplesmente conexas) define a classe de genus $g = 0$, seguido do toro e seus congêneres (xícaras, seres humanos, etc.) com $g = 1$, do bi-toro ($g = 2$, uma superfície com duas alças), etc. Veja a Figura (1.2) para uma ilustração destes genii e homeomorfismos. Na verdade, pode-se mostrar que as variedades bidimensionais orientáveis são sempre homeomorfas a uma das variedades de curvatura constante, com respectivamente, curvatura positiva, nula ou negativa: a 2–esfera ($g = 0$), o plano ($g = 1$, como veremos a seguir), ou a sela ($g \geq 2$).

De forma mais relevante para a topologia cósmica, variedades tridimensionais de curvatura constante positiva e nula também foram totalmente classificadas, mas as de curvatura negativa ainda desafiam uma categorização completa [169].

Considere um toro bidimensional, ou 2–toro, definido geometricamente como a superfície de revolução gerada por uma circunferência que não intersepta o eixo de revolução. Este é um objeto que pode ser visualizado em um espaço tridimensional.

Pela descrição acima é fácil obter a métrica desta superfície a partir da métrica euclidiana de $\mathbb{R}^3$

$$ds^2 = (R + r \cos \phi)^2 (d\theta^2 + r^2 d\phi^2) , \quad (1.30)$$

onde r é o raio do círculo e R a distância de seu centro ao eixo de revolução. A curvatura gaussiana é dada neste caso por

$$K = \frac{\cos \phi}{r(R + r \cos \phi)} . \quad (1.31)$$

Como é obvio tanto a partir da Figura 1.3 quanto da fórmula acima, o toro assim definido possui regiões com curvatura positiva, negativa e nula. Mas considere dois cortes circulares, um ao longo do equador perpendicular ao eixo de simetria, e outro seccionando ao longo do círculo gerador. Topologicamente, é possível recuperar o toro completo esti-

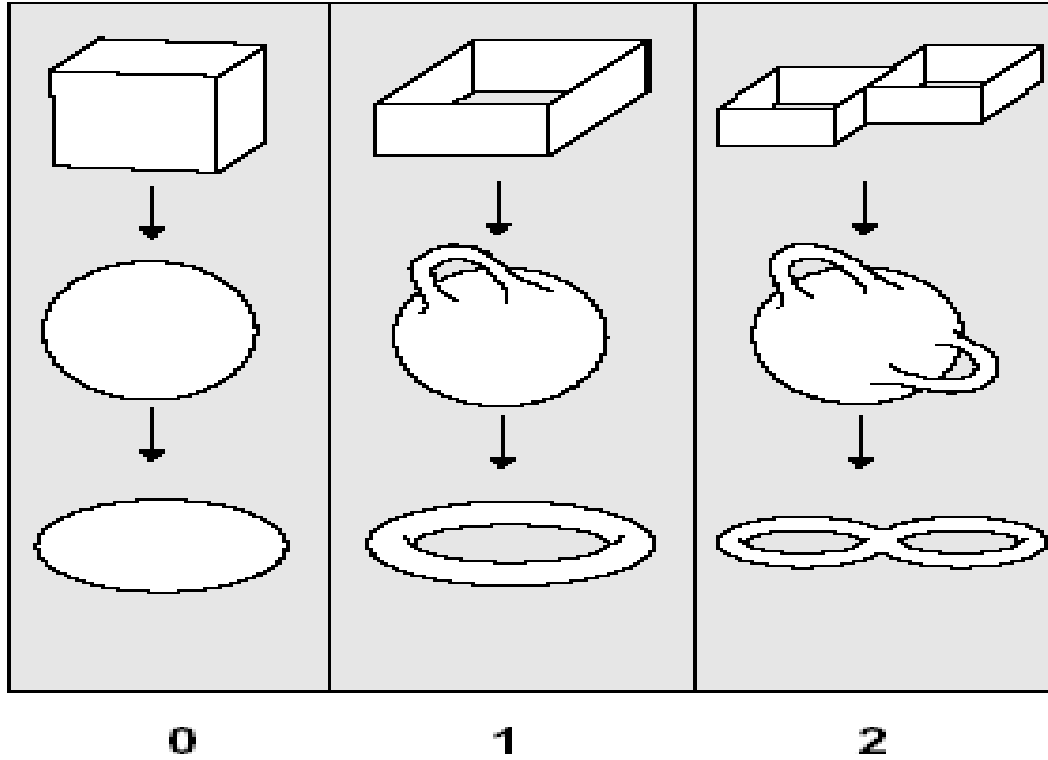


Figura 1.2: As classes de homeomorfismos para superfícies compactas orientáveis, com alguns exemplos simples. O número denota o genus, ou número de buracos, de cada classe. *Fonte: M. Lachieze-Rey & J.-P. Luminet (1999)*

cando a superfície resultante na forma de um retângulo e identificando formalmente seus lados opostos (novamente, veja a Figura 1.3). Intuitivamente, um ponto próximo de uma das arestas está próximo de todas as demais arestas; um hipotético viajante que saísse pelo lado direito emergiria imediatamente do lado esquerdo. É possível mostrar que variedades topológicas em duas ou três dimensões podem sempre ser representadas desta forma: um domínio fundamental (o retângulo, neste caso; mais geralmente um polígono em duas dimensões e um poliedro em três), no qual pares de arestas (faces, em três dimensões) são identificadas entre si. Conjuntamente, o domínio fundamental e as regras de identificação (ou isometrias) determinam a topologia.

Definido desta forma, o toro não tem a priori uma métrica característica. Com uma escolha apropriada de sistema de coordenadas, ele poderia ser dotado da métrica (1.30), por exemplo. Mas suponha que, assim como fazemos em relação à topologia do Universo, exigimos ainda que o toro seja dotado de uma métrica com curvatura constante, que pode ser sempre escrita como

$$ds^2 = a^2(d\chi^2 + f^2(\chi) d\phi^2) . \quad (1.32)$$

Assim como no caso tridimensional, em duas dimensões existem somente três opções

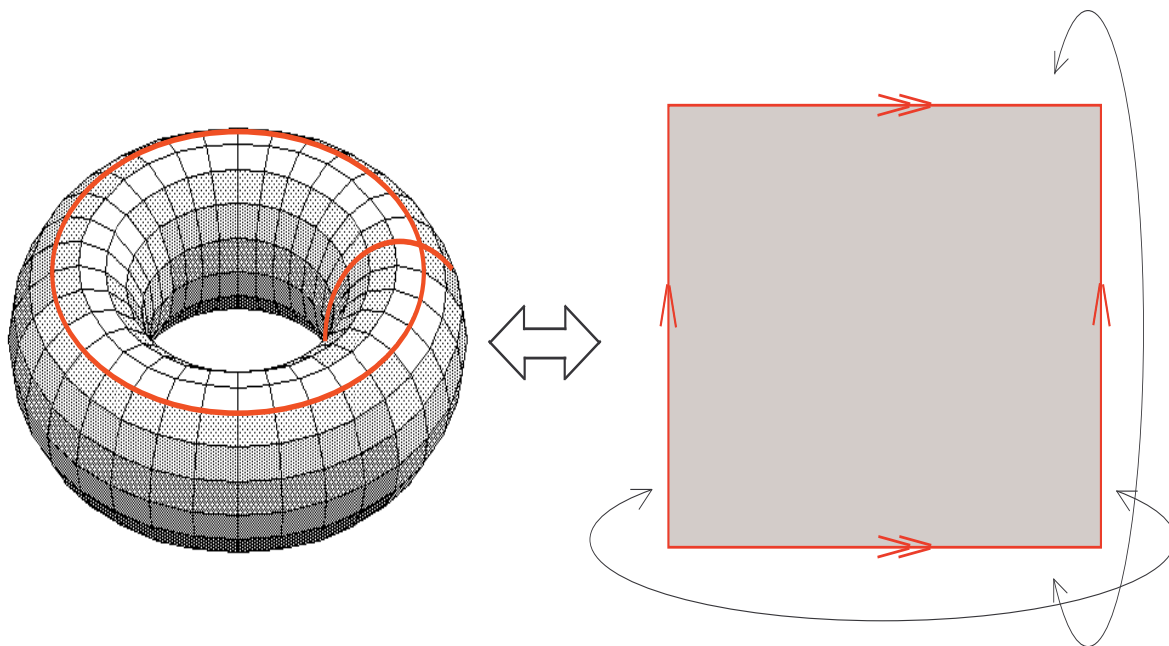


Figura 1.3: O 2–toro pode ser representado geometricamente em $\mathbb{R}^3$, ou através da identificação das arestas do seu domínio fundamental em $\mathbb{R}^2$. As setas indicam quais arestas devem ser identificadas (ou ‘coladas’), e como. Diferentes regras geram diferentes topologias.

possíveis para o espaço de cobertura com curvatura constante (cujo valor é dado, em termos fator de escala a na Equação (1.32), por $K = k/a^2$): um plano ($k = 0$), uma esfera ($k = 1$) ou uma superfície em forma de sela ($k = -1$), onde $f(\chi)$ é respectivamente χ , $\sin \chi$ e $\sinh \chi$. Mas note que os vértices do domínio fundamental são todos imagens do mesmo ponto. Uma circunferência centrada neste ponto pode ser então subdividida em 4 quadrantes, cuja soma de ângulos é obviamente 2π . Estes quatro ângulos também correspondem aos ângulos internos do retângulo definido pelo domínio fundamental. Assim, a somatória dos ângulos internos de um quadrângulo em um toro, que indicamos por $S_{\square}$, é exatamente 2π ! Em geral, em superfícies de curvatura constante, $S_{\square} = 2\pi + KA$, onde A é a área do quadrângulo. Assim, no caso do 2–toro de curvatura constante temos que $K = 0$, e portanto a curvatura é necessariamente nula. Note ainda que esta propriedade não depende de como os lados do domínio fundamental são identificados, o que implica que ela se aplica a qualquer variedade topológica em duas dimensões com um domínio fundamental quadrangular.

Lema 1.1 *Se um 2–toro (ou qualquer outra variedade topológica em duas dimensões com um domínio fundamental retangular) tem curvatura gaussiana constante, então a curvatura é identicamente nula.*

Generalizando, é fácil ver que um resultado similar vale para qualquer domínio fun-

damental que ladrilhe perfeitamente o espaço euclidiano. Um, hexágono com as faces opostas identificadas, por exemplo, é o domínio fundamental de um toro com geradores não-ortogonais. Em três dimensões, 3-toros e outras 3-variedades de curvatura constante cujo domínio fundamental é um paralelepípedo reto (e.g., um cubo) ou um prisma de base hexagonal têm forçosamente curvatura identicamente nula.

O Lema 1.1, e a sua extensão para 3-variedades, ilustra a relação entre geometria e topologia. Ao restringir a geometria, restringimos a classe de topologias possíveis, e vice versa. Assim, se as seções espaciais do Universo tiverem a topologia de um 3-toro, então o Universo tem seção espacial euclidiana. De forma análoga, se o Universo não for plano então suas seções espaciais não são homeomorfas a um toro. Temos então, sem muito esforço, um resultado que é o protótipo da classificação das topologias possíveis para as seções espaciais do Universo. Estudaremos no Apêndice B as diferentes classes de topologias compatíveis com seções espaciais do Universos localmente planas, esféricas e hiperbólicas.

Podemos ainda considerar o ponto de vista de um hipotético observador confinado ao 2-toro definido por um domínio fundamental quadrado de lado unitário. Suponha que nosso intrépido amigo bidimensional tivesse uma vela acesa nas mãos. Ele certamente poderia ver a vela que carrega diretamente; mas além disso, se olhasse em uma das direções perpendiculares a uma das faces que delimitam o domínio fundamental, veria também outra imagem da mesma vela, situada a uma distância igual a 1. De fato, ele veria um mosaico de imagens repetidas da vela até onde a sua vista alcançasse, distribuídas nos vértices de um gride de lado unitário. Cada imagem corresponde a um caminho possível que a luz emitida pela vela pode percorrer para voltar ao lugar de origem. Nosso observador poderia então supor talvez que seu 'Universo' fosse dotado de uma geometria muito peculiar, capaz de distorcer a trajetória da luz emitida pela vela de modo a formar exatamente o mosaico observado. Mas se houvesse razão física para supor que a trajetória dos raios de luz deveria ser retilínea, ele poderia então usar a parte do mosaico que ele é capaz de ver para tentar deduzir a topologia do espaço em que vive. Se só visse a vela olhando diretamente para as próprias mãos, seria forçado a concluir que habita um espaço onde não existem caminhos fechados para raios de luz, ou então que a sua vela não é luminosa o suficiente para que mesmo a sua imagem mais próxima possa ser vista com clareza. Em suma, se quisesse restringir as topologia possíveis do seu Universo a partir de observações, nosso habitante do toro teria primeiramente que considerar as condições necessárias para sua detectabilidade.

Para concluir, mostramos um segundo exemplo de variedade topológica bidimensional, denominada garrafa de Klein, que tem o mesmo domínio fundamental que o toro mostrado

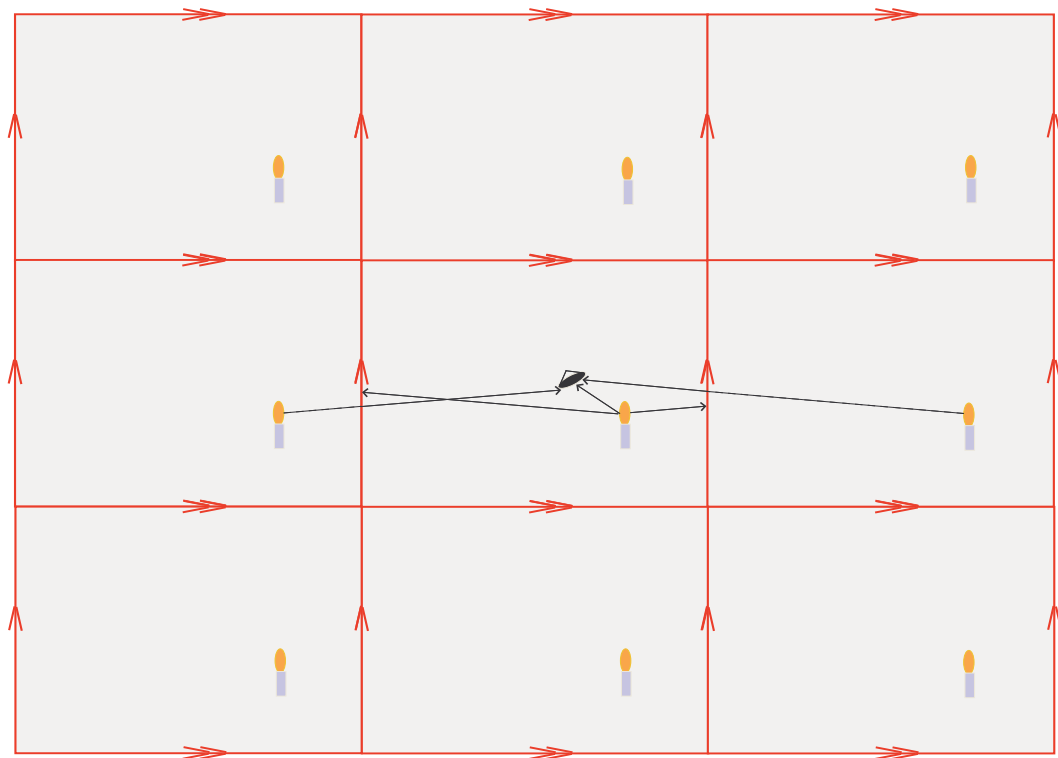


Figura 1.4: As múltiplas imagens de uma fonte luminosa pontual vista por um observador em um toro. Três das trajetórias fechadas da luz emitida são indicadas (linhas sólidas) dentro do domínio fundamental. O padrão assim obtido é igual (linhas tracejadas) ao que seria gerado por um mosaico no espaço de recobrimento, composto de cópias idênticas do domínio fundamental, unidas de forma a respeitar as regras de identificação das arestas. Note que as dimensões do domínio fundamental determinam as escalas do mosaico, mas que este não depende da escolha do ponto-base.

acima, mas com diferentes regras de identificação das arestas (na Figura 1.5, as arestas laterais são identificadas após uma rotação de π). O mosaico resultante é notadamente diferente do mosaico correspondente ao toro. Se fosse capaz de observar as primeiras imagens da vela à direita e à esquerda, o nosso observador poderia distinguir entre as duas topologias. Note que, ao contrário do toro, na garrafa de Klein o padrão do mosaico depende da posição da 'vela' no domínio fundamental.

1.4.3 Homotopias por laço, holonomias e espaço de recobrimento

No exemplo do 2-toro discutido acima, sugerimos que os caminhos fechados percorridos pela luz emitida por uma vela podem ser usados para diferenciar entre topologias. Podemos imaginar cada um destes caminhos fechados como um laço, que começa e termina em

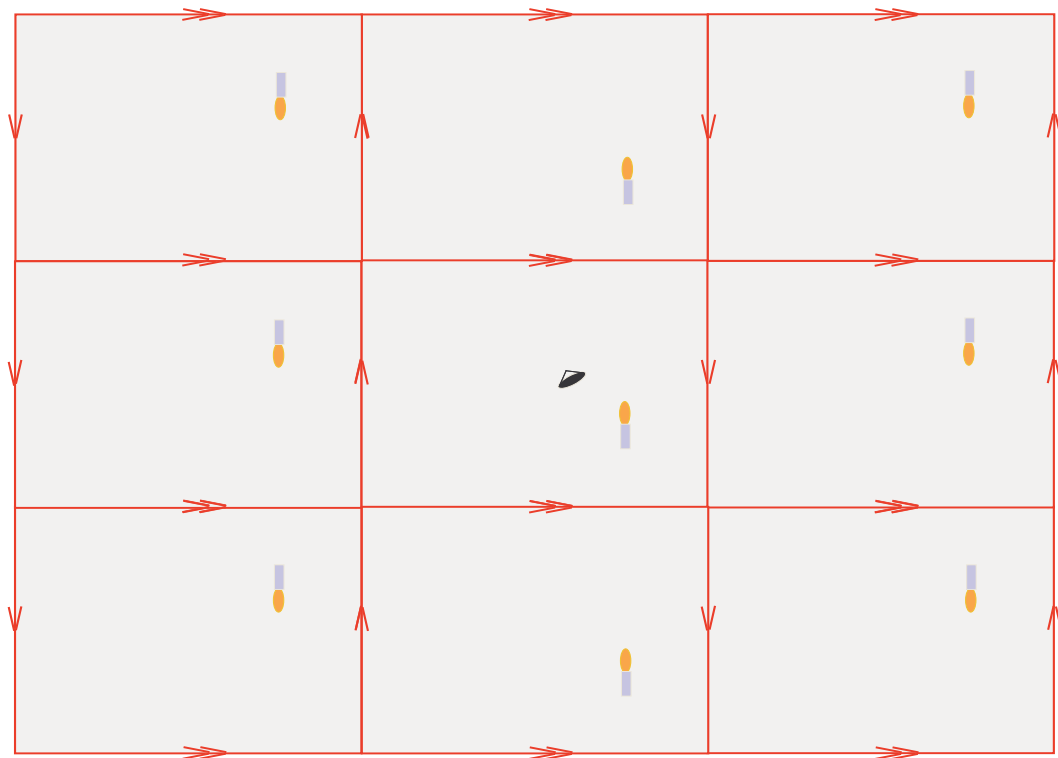


Figura 1.5: Mosaico obtido a partir de uma regra diferente de identificação das arestas do domínio fundamental. Note que, ao contrário do toro, nesta dita garrafa de Klein o padrão do mosaico depende da escolha do ponto-base.

um ponto de origem. Ainda usando a intuição do toro, note que cada uma das imagens da vela define um laço que não pode ser deformado continuamente em nenhum dos outros; eles podem ser caracterizados pelo número de vezes que a luz emitida perfaz uma volta completa no domínio fundamental em cada um dos dois sentidos definidos pelas isometrias que identificam as arestas. Existe portanto uma correspondência um-a-um entre as classes de equivalência por deformação contínua dos laços em um toro, e os pares ordenados do conjunto $(\mathbb{Z}, \mathbb{Z})$. Qualquer laço que não cruze as arestas do domínio fundamental pode ser deformado continuamente em um ponto, e corresponde assim ao par $(0, 0)$. Além disso, a operação de 'emendar' dois laços que passam pelo mesmo ponto pode ser vista como formalmente equivalente a uma operação de grupo, com uma estrutura isomorfa ao grupo formado pelos pares ordenados de inteiros, $(\mathbb{Z}, \mathbb{Z})$, dotados de uma operação de soma. Intuitivamente, parece mais simples estudar a estrutura deste grupo simples do que o conjunto das variedades homeomorfas ao toro.

De forma mais geral, esta estrutura de grupo dos laços pode ser usada para classificar diferentes topologias de uma forma mais simples do que os homeomorfismos; mas embora sejam suficientes para classificar variedades bidimensionais compactas e orientáveis, as classes de laços não fornecem em geral uma classificação completa (no sentido de que

algumas topologias podem ser equivalentes por laços mas não homeomorfas entre si). Para captar algumas estruturas topológicas mais complexas, é possível definir o equivalente a 'laços' em dimensões mais altas (essencialmente, laços são deformáveis em círculos, ou 1-esferas; estruturas de ordens mais altas são deformáveis em n -esferas). Isto motiva a definição

Definição 1.8 Uma **homotopia** existe entre duas funções contínuas f e g que levam o espaço topológico M no espaço topológico W se existe uma função contínua $H : M \times [0, 1] \rightarrow W$ tal que para qualquer $x \in M$, $H(x, 0) = f(x)$ e $H(x, 1) = g(x)$. Os espaços topológicos M e W são equivalentes por homotopia se existirem funções f e f^{-1} tais que haja uma homotopia entre $f \circ f^{-1}$ e a identidade I .

Podemos imaginar o parâmetro t de $H(x, t)$ como um tempo, ao longo do qual a função f é deformada suavemente em g . Note que segundo a definição dois espaços homeomorfos são sempre equivalentes por homotopia (H , f e g são neste caso bijeções bicontínuas), mas o contrário não é necessariamente verdade.

Em geral, as classes de homotopia mais úteis em topologia algébrica são as que relacionam n -esferas com espaços topológicos a serem estudados. Grosso modo, estas classes traduzem as propriedades topológicas essenciais na proximidade de cada ponto em termos de estruturas mais simples, como laços, bolas e suas generalizações de dimensões superiores. Formalmente

Definição 1.9 Para um espaço M com um **ponto-base** x , $\pi_n(M, x)$ é o conjunto das classes de equivalência por homotopia das funções $\alpha : S^n \rightarrow M$ para as quais $\alpha(y) = x$, onde y é o ponto base de S^n . É possível mostrar que os elementos de π_n formam um grupo para cada $n \geq 1$, e π_n é portanto denominado **grupo de homotopia**.

Assim, em cada ponto de uma variedade topológica, as classes de homotopia definem famílias de n -superfícies fechadas deformáveis entre si, e que mantêm o ponto-base invariante. Para os nossos propósitos, só precisaremos estudar o primeiro grupo de homotopia π_1 , também conhecido como grupo fundamental (c.f. [95] para uma abordagem exaustiva e rigorosa dos conceitos apresentados nesta seção e na seguinte).

Para definir a operação de grupo de π_1 , é possível entender a definição acima, restrita ao grupo fundamental, em termos de caminhos fechados, ou laços. Assim, para um espaço M e um ponto base x , um **laço** é definido como a imagem de uma função contínua $\alpha : [0, 1] \rightarrow M$ tal que $\alpha(0) = \alpha(1) = x$.⁵ Dois laços são homeotópicos quando compartilham

⁵Abusando ligeiramente da linguagem, diremos simplesmente 'laço α ', ao invés do trava-línguas 'imagem do intervalo unitário pela função contínua α '.

o mesmo ponto-base, e quando as funções das quais são imagens são equivalentes por homotopia. Definidos desta forma, é fácil entender a operação de produto entre laços como a formalização da noção intuitiva de ligá-los por seu ponto-base. Assim,

Definição 1.10 *Sejam α e β laços em M com ponto-base comum x . O **produto de laços** $h = \beta \circ \alpha$ é definido como $\gamma : [0, 1] \rightarrow M$; $\gamma(t) = \beta(2t)$ se $0 \leq t \leq 1/2$ e $\gamma(t) = \alpha(2t - 1)$ se $1/2 \leq t \leq 1$.*

Definido desta forma, h é claramente também um laço com ponto base x , no qual os laços constituintes β e α são percorridos sucessivamente, com o dobro da 'velocidade' paramétrica. Existe obviamente uma definição mais geral para operações de grupo de homotopias de ordem superior, que não vem ao caso aqui.

Considere agora em uma variedade topológica M o conjunto de todas os laços α' que são homeotópicos a α (denotaremos esta relação doravante como $\alpha \sim \alpha'$). Esta classe de equivalência, que denotamos por $[\alpha]$, é um elemento do grupo fundamental de M . Para que o produto entre laços possa ser usado como a operação de grupo de $\pi_1(M, x)$, é preciso que os laços resultantes do produto entre todos os elementos de duas classes de laços sejam homeotópicos entre si. Ou seja,

Lema 1.2 *Se $\alpha' \in [\alpha]$ e $\beta' \in [\beta]$, e $\alpha \circ \beta = \gamma$, então $\alpha' \circ \beta' = \gamma' \in [\gamma]$.*

Prova: A demonstração é obtida diretamente a partir da definição de homotopia. Basta provar que existe uma função contínua $H : [0, 1] \times [0, 1] \rightarrow M$ tal que deforma suavemente γ em γ' . (i.e., para o ponto-base $x \in M$, $H(0, x) = H(t, x)$, $H(0, *) = H(1, *)$, e $H(t, 0) = \gamma(t)$ e $H(t, 1) = \gamma'(t)$). Mas por definição existem funções contínuas H_α e H_β que deformam respectivamente α em α' e β em β' . Assim, de forma análoga ao produto entre laços, definimos $H = H_\alpha(2t)$ se $0 \leq t \leq 1/2$ e $H = H_\beta(2t - 1)$ se $1/2 \leq t \leq 1$. Desta maneira, por construção $\gamma' \in [\gamma]$ ■

Com base no Lema 1.2, definimos o produto entre classes de laços da seguinte maneira: Se $\alpha \circ \beta = \gamma$, então $[\alpha] \circ [\beta] = [\gamma]$, que formalmente pode ser lido como 'o produto de um elemento de $[\alpha]$ por um elemento de $[\beta]$ é um elemento de $[\gamma]$ '.

Note que em um espaço conexo M , a escolha do ponto-base não muda a classe de equivalência de um laço. De fato, quaisquer dois pontos x e x' podem ser ligados por um laço $\nu_{xx'} : [0, 1] \rightarrow M$ que vai e volta pelo mesmo caminho (i.e. $\nu_{xx'}(0) = x$, $\nu_{xx'}(1/2) = x'$ e $\nu_{xx'}(x) = \nu_{xx'}(1 - x)$). Um tal laço, dito trivial, pode ser sempre contraído continuamente a um ponto.

Lema 1.3 *Em de um espaço conexo M , os conjuntos das classes de equivalência por homotopia de laços, $\pi_1(M)$ para quaisquer ponto-bases $x \in M$ e $x' \in M$ são isomorfos*

Prova: A definição de $\nu_{xx'}$, acima, implica que, se $\alpha \sim \alpha'$ com ponto-base x , então $\alpha \circ \nu_{xx'} \sim \alpha' \circ \nu_{xx'}$ com ponto-base x' ■

Portanto, a menos de um isomorfismo, podemos denotar o grupo fundamental $\pi_1(M, x)$ de um espaço conexo como $\pi_1(M)$ sem ambiguidade. Finalmente, de posse destes lemas e definições, podemos provar que

Teorema 1.1 *O conjunto das classes de equivalência por homotopia de laços, $\pi_1(M)$, forma um grupo, fechado por operação de produto de laços.*

Prova: Provamos diretamente as quatro propriedades que definem um grupo:

- 1 - Fechamento: Pelo Lema 1.2, se $[\alpha]$ e $[\beta] \in \pi_1$, então $[\alpha \circ \beta] \in \pi_1$
- 2 - Associatividade: Pela definição do produto de laços, $\alpha \circ (\beta \circ \gamma) \sim (\alpha \circ \beta) \circ \gamma$. Pelo Lema 1.2, $[\alpha] \circ ([\beta] \circ [\gamma]) = ([\alpha] \circ [\beta]) \circ [\gamma]$
- 3 - Elemento identidade: Para qualquer elemento $\nu_{xx'}$ da classe dos laços deformáveis em um ponto, $[\nu]$, $\nu_{xx'} \circ \alpha \sim \alpha \sim \alpha \circ \nu_{xx'}$. Assim, pelo Lema 1.2 $[\nu] \circ [\alpha] = [\alpha] = [\alpha] \circ [\nu]$
- 4 - Elemento inverso: Seja $\gamma : [0, 1] \rightarrow M$ um laço. Então o laço definido por $\gamma^{-1}(t) = \gamma(1 - t)$ é tal que o produto $\gamma \circ \gamma^{-1}$ vai e volta pelo mesmo caminho, e é assim um laço trivial $\nu_{xx} \in [\nu]$. Portanto, pelo Lema 1.2, $[\gamma] \circ [\gamma^{-1}] = [\nu] = [\gamma^{-1}] \circ [\gamma]$ ■

Um espaço topológico no qual o grupo fundamental consiste somente do elemento identidade (ou seja, no qual todos os laços são deformáveis uns nos outros) é chamado **simplesmente conexo**. Como vimos, dois espaços homeomorfos são sempre equivalentes por homotopia, e portanto o grupo fundamental definido acima é um invariante topológico, no sentido de que espaços homeomorfos possuem sempre o mesmo grupo fundamental. A recíproca não é necessariamente verdadeira; existem espaços dotados do mesmo grupo fundamental que não são homeomorfos.⁶

O último passo nesta seção é relacionar o grupo fundamental de uma variedade multiplamente conexa a um grupo isomorfo de isometrias definidas em uma variedade sim-

⁶Henri Poincaré [134] conjecturou em 1904 que qualquer variedade compacta n -dimensional com um grupo fundamental trivial (i.e. onde todos os laços podem ser contraídos a um ponto) deveria ser homeomorfa à n -esfera. Ao longo dos anos a conjectura foi provada para todas as dimensões exceto $n = 3$. Este último caso se mostrou renitente até 2006, quando foi demonstrado por Grigory Perelman, um matemático russo recluso. Após uma busca frenética, conduzida por matemáticos e jornalistas do mundo inteiro, Perelman foi localizado morando na casa de sua mãe, em São Petersburgo, mas humildemente recusou qualquer prêmio ou recompensa pelo feito, incluindo a medalha Fields em agosto de 2006.

plesmente conexa. Em topologia cósmica, a forma global do espaço pode ser expressa de maneira puramente geométrica em termos de tais isometrias.

Na seção anterior, vimos como os 'caminhos fechados' distintos de um toro (i.e., os elementos do seu grupo fundamental, para usar os conceitos mais precisos definidos acima), correspondem cada um a uma cópia do domínio fundamental que, juntas, cobrem um espaço simplesmente conexo, dito de recobrimento universal. Este é o protótipo do processo que pode ser estendido de forma unívoca para todas as variedades localmente conexas, homogêneas e isotrópicas. Portanto

Definição 1.11 *Um **recobrimento** de um espaço topológico M é um espaço C para o qual exista um mapa contínuo sobrejetivo $f : C \rightarrow M$ tal que $\forall x \in M$ exista uma vizinhança V de x onde $f^{-1}(V) = \bigcup_i U_i$, onde U_i são abertos disjuntos, $U_i \cap U_j = \emptyset$ se $i \neq j$, e $f|_{U_i} : U_i \rightarrow V$ é um homeomorfismo local sobrejetivo. Um espaço de recobrimento é dito **universal** se é simplesmente conexo. O conjunto dos pontos em C cuja imagem em M é x é denominado **órbita** de x .*

Claramente, nenhum laço não trivial presente no espaço-base é preservado quando o espaço de recobrimento é universal, já que o grupo fundamental deste último é trivial. É possível ainda mostrar que para toda variedade topológica conexa o espaço de recobrimento universal é único, a menos de homeomorfismos. Para maiores detalhes a respeito de espaços de recobrimento e operações de levantamento, no contexto da topologia cósmica, recomendo a tese de doutorado de Germán Gomero [62].

Intuitivamente, um espaço de recobrimento consiste de um mosaico de cópias localmente idênticas do espaço-base. Por exemplo, o círculo unitário S^1 pode ser parametrizado em $\mathbb{R}^2$ como $c(\theta) = (\cos \theta, \sin \theta)$. A reta $\mathbb{R}$ é portanto o espaço de recobrimento universal de S^1 , através do mapa $c : \mathbb{R} \rightarrow S^1$. Cada ponto do círculo é representado por infinitos pontos na reta, separados por intervalos de 2π . Note que um laço de ordem n com ponto-base x no círculo corresponde a um segmento de comprimento $2\pi n$ na reta, ligando dois pontos na órbita de x . De forma mais geral, é possível definir as transformações que relacionam classes de equivalência de laços no espaço-base com órbitas do ponto-base no espaço de recobrimento.

Definição 1.12 *Um **automorfismo** de um recobrimento $f : \widetilde{M} \rightarrow M$ é um homeomorfismo local $g : \widetilde{M} \rightarrow \widetilde{M}$ tal que $f \circ g = f$. O conjunto dos automorfismos forma um grupo denotado $\text{Aut}(f, \widetilde{M}, M)$*

Algumas propriedades de $\text{Aut}(f, \widetilde{M}, M)$ podem ser obtidas diretamente a partir da definição acima. Primeiramente, $\text{Aut}(f, \widetilde{M}, M)$ é um grupo discreto, porque π_1 é discreto.

Além disso, a ação de $\text{Aut}(f, \widetilde{M}, M)$ é **livre** (sem pontos fixos exceto para a identidade) e **propriamente descontínua** (i.e., $\forall y \in \widetilde{M}$ existe sempre uma vizinhança U tal que $g(U) \cap U = \emptyset$, novamente excluindo a identidade). As duas últimas propriedades são consequência da definição de um espaço de recobrimento, que garante a existência de vizinhanças disjuntas em $\widetilde{M}$ que contém cada uma somente um elemento da órbita de $x = f^{-1}(y)$, e do seguinte lema:

Lema 1.4 *Dado qualquer $x \in M$ e um ponto qualquer de sua órbita $y \mid f(y) = x$, a ação de $\text{Aut}(f, \widetilde{M}, M)$ sobre y gera a órbita de x .*

Prova: Se y e y' pertencem à órbita de x , então pela definição de recobrimento estes pontos possuem vizinhanças V_y e $V_{y'}$ que são homeomorfas entre si. Existe assim um homeomorfismo local $g : V_y \rightarrow V_{y'}$ tal que $g(y) = y'$ e $f(g(w)) = f(w) \forall w$, e portanto $g \in \text{Aut}(f, \widetilde{M}, M)$.

Finalmente, é possível provar (veja [62], seção 1.4.2, para uma demonstração) que

Teorema 1.2 *Se $f : C \rightarrow M$ define um recobrimento universal, então o grupo de automorfismos de $\widetilde{M}$ é isomorfo ao grupo fundamental de M , i.e., $\text{Aut}(f, \widetilde{M}, M) \approx \pi_1(M)$.*

Como consequência, a estrutura topológica contida no grupo fundamental é preservada nos automorfismos. Veremos a seguir que no caso que nos interessa, das 3-variedades com curvatura constante, os automorfismos equivalem a isometrias, que são as estruturas efetivamente detectáveis em topologia cósmica.

1.4.4 Domínio fundamental, e as isometrias da topologia cósmica

Retomando nossa pergunta original: Qual a forma do Universo? A resposta completa envolve tanto propriedades geométricas, definidas localmente em cada ponto, quanto características topológicas, que se referem ao espaço como um todo. Geometricamente, é bem entendido que o Universo é bem descrito por uma variedade pseudo-riemanianna dotada de uma métrica FLRW. Como vimos, existe uma escolha natural de um tempo universal, definido pelo tempo próprio de observadores em repouso em relação ao fluxo de Hubble, que é sempre perpendicular às seções espaciais de curvatura espacial constante. Assim, a métrica (1.10) pode ser escrita simplesmente

$$g = -dt \otimes dt + a(t)^2 h, \quad (1.33)$$

onde h é a métrica de uma 3-variedade riemanianna conexa M , localmente homogênea e isotrópica, e portanto de curvatura constante. Neste caso, qualquer que seja a topologia

de M , é possível definir univocamente um espaço de cobertura universal $\widetilde{M}$ simplesmente conexo. Mas localmente a geometria do espaço não varia se este é multiplamente conexo ou não. Assim, o espaço de cobertura deverá ser dotado da mesma métrica e do mesmo grupo fundamental que uma das 3-variedades simplesmente conexas de curvatura constante possíveis.

Considere agora o grupo das isometrias de $\widetilde{M}$, $I(\widetilde{M})$, o grupo das bijeções $\widetilde{M} \rightarrow \widetilde{M}$ que preservam distâncias e ângulos. Este grupo será isomorfo a $I(\mathbb{R}^3)$, $I(\mathbb{S}^3)$ ou $I(\mathbb{H}^3)$. Mas espaços homogêneos são por definição aqueles sobre os quais suas isometrias agem transitivamente, i.e., se $x, y \in \widetilde{M}$, existe sempre uma isometria $\phi \in I(\widetilde{M})$ para a qual $\phi(x) = y$. Quaisquer dois pontos na variedade são ligados por pelo menos uma isometria. Em uma variedade riemanianna geodesicamente completa, cada uma destas isometrias corresponde a (pelo menos) uma geodésica.

Seja portanto y um ponto-base em $\widetilde{M}$. A ação de cada elemento $g \in \text{Aut}(\widetilde{M}, M)$ ⁷ do grupo de automorfismos sobre y gera um ponto $g(y)$ da órbita de y ; mas em espaços de curvatura constante sem fronteira (que são os que nos interessam), existe sempre uma isometria γ tal que $\gamma(y) = g(y)$. É possível mostrar que para todo automorfismo g existe uma e somente uma isometria γ que tem a mesma ação sobre todos os pontos do espaço, e que o conjunto destes γ forma um subgrupo $\Gamma_M \subset I(\widetilde{M})$ isomorfo a $\text{Aut}(\widetilde{M}, M)$.

O grupo Γ_M é denominado grupo de **holonomia** de M em $\widetilde{M}$. Devido às propriedades dos automorfismos discutidas acima, ele é discreto, e de ação livre e propriamente descontinua. Estas propriedades em particular garantem a existência, para cada ponto-base x , de uma vizinhança U na qual $\gamma(U) \cap U = \emptyset$ para todo $\gamma \in \Gamma_M$. A maior vizinhança conexa de x que possui esta propriedade é denominada **domínio fundamental**. O domínio fundamental tem o mesmo volume que M .

Pela definição do espaço de recobrimento universal, a ação do grupo Γ_M cobre completamente e sem sobreposições o espaço de cobertura universal com cópias do domínio fundamental. Cada ponto no fecho do domínio fundamental (a sua 'fronteira') é levado por um elemento de Γ_M em algum outro ponto também no fecho. Como Γ_M é discreto, o número de tais isometrias é finito, e portanto o fecho do domínio fundamental consiste um número finito de faces, duas a duas idênticas (por serem identificadas por uma isometria) de um poliedro⁸ convexo (necessário para que o espaço de cobertura possa ser tesselado). Este é o chamado **poliedro fundamental**.

Embora tenhamos definido o espaço de recobrimento universal a partir do espaço-

⁷Aqui não especificamos o mapa entre M e $\widetilde{M}$ por ser ele univocamente determinado.

⁸Para um espaço de cobertura tridimensional. Em duas dimensões teríamos um polígono, como no caso do 2-toro. Obviamente o domínio fundamental tem a mesma dimensão que o espaço de cobertura.

base, é possível fazer o caminho inverso. Dada uma variedade Σ , e um grupo discreto Γ de isometrias de ação livre e propriamente descontínua que age sobre Σ , a variedade quociente Σ/Γ é definida como o conjunto das órbitas, ou classes de equivalência, que existem em Σ pela ação de Γ ,

$$\frac{\Sigma}{\Gamma} \equiv \{O_{\Sigma}(y) \equiv [y] : y \in \Sigma\} , \quad (1.34)$$

onde $O_{\Sigma}(y)$ é a órbita de y pelo grupo Γ . Existe uma definição mais geral, puramente em termos de ação de grupos, mas esta nos basta.

O mapa $\phi : \Sigma \rightarrow \Sigma/\Gamma$ que leva $\phi(y) \rightarrow O_{\Sigma}(y)$ é a projeção de Σ no espaço das órbitas O_{Σ} . As propriedades exigidas de Γ garantem que ϕ é uma aplicação de recobrimento, e que Σ/Γ também é uma variedade topológica. Matematicamente, este é exatamente o processo inverso da obtenção de um espaço de cobertura; assim, claramente, podemos escrever $M = \widetilde{M}/\Gamma_M$. A obtenção de um espaço-quociente preserva a dimensão e propriedades geométricas locais do espaço original; mas não preserva topologia. É portanto uma maneira de gerar espaços multiplamente conexos, possivelmente compactos, a partir de um espaço simplesmente conexo. De fato, é possível provar [25] que todas as 3-variedades sem fronteiras de curvatura seccional constante podem ser geradas desta forma a partir de alguma variedade simplesmente conexa.

Lembrando que o grupo fundamental de uma variedade topológica conexa e o grupo de holonomias são isomorfos, podemos portanto (em princípio) obter as topologias compatíveis com um espaço de recobrimento universal Σ gerando os espaços-quociente de todos os subgrupos Γ_i de $I(\Sigma)$ que sejam discretos e de ação livre e propriamente descontínua (infelizmente, para o caso hiperbólico, não existe uma classificação completa de tais grupos, e nem, portanto, das topologias). O domínio fundamental pode ser obtido pela aplicação sistemática dos elementos do grupo Γ a um ponto-base, embora na prática só um subconjunto completo bem escolhido de seus geradores precisa ser levado em conta.

Dois tipos de isometria são de particular importância: As **translações de Clifford** são isometrias que movem todos os pontos da variedade por uma mesma distância (formalmente, se γ é uma translação de Clifford, para quaisquer pontos u e w na variedade, temos que $d(u, \gamma u) = d(w, \gamma w)$). O conjunto das translações de Clifford forma um subgrupo abeliano, pois comutam entre si, e também comutam com as demais translações de $I(\widetilde{M})$ [124]. Quando não houver possibilidade de confusão, vamos nos referir a tais isometrias somente como 'translações'. Outro conjunto importante de isometrias, definido para cada ponto x da variedade, é o chamado **grupo de isotropias** em x , que consiste em todas as isometrias que tem x como ponto fixo, e coincide com a noção intuitiva de rotações

(próprias e impróprias) em torno de x . Uma 3-variedade é **globalmente isotrópica** se possui para cada ponto um grupo de isotropias dotado de uma representação que inclui como subgrupo $SO(3)$, o grupo das rotações em $\mathbb{R}^3$.

O grupo de isometrias do espaço quociente $M = \widetilde{M}/\Gamma_M$ é homomorfo (i.e., preserva a estrutura) ao subgrupo de $I(\widetilde{M})$ que comuta⁹ com Γ_M , o chamado **normalizante** de Γ em $I(\widetilde{M})$, $N(\Gamma)$. Assim, $I(M) = N(\Gamma)/\Gamma$ (c.f. o artigo pioneiro de Ellis, [50]). Em termos intuitivos, $I(M)$ consiste nas operações que levam o domínio fundamental em uma cópia idêntica de si mesmo. Particularmente, em um espaço dito **globalmente homogêneo**, Γ_M contém somente translações de Clifford, que devido às suas propriedades de comutação implicam que o domínio fundamental é o mesmo para qualquer ponto-base (ou, fisicamente, observador); neste caso as únicas simetrias quebradas pela obtenção do quociente são rotacionais. Note que as propriedades espaciais de isotropia e homogeneidade globais descritas acima são mais restritivas do que as hipóteses de isotropia e homogeneidade locais que implicam na métrica de FLRW (1.10), que só exigem invariância por translações e rotações de uma vizinhança de cada ponto, não do espaço inteiro.

Dizemos ainda que uma 3-variedade $M = \widetilde{M}/\Gamma$ é **orientável** se em qualquer ponto toda isotropia γ de M pode ser expressa como a composição $-I \circ \gamma'$, onde $-I$ denota uma inversão (na representação matricial de $O(3)$, $-I = \text{diag}\{-1, -1, -1\}$) e $\gamma' \in SO(3)$, também é isotropia de M . Esta definição corresponde à noção intuitiva de uma variedade na qual 'direita' e 'esquerda' podem ser definidas univocamente em qualquer caminho fechado.

O conjunto das 3-variedades orientáveis de curvatura constante é infinito, e quase infinitamente variado. Esboçamos a sua classificação no apêndice B. Maiores detalhes podem ser encontrados em [189], [62], [142] e [143].

Finalmente, para fixar os conceitos que desenvolvemos após tanta labuta, que são fundamentais no estudo da topologia cósmica, relacionamo-os com as noções intuitivas desenvolvidas na Seção 1.4.2. No exemplo do 2-toro, o domínio fundamental é um retângulo, cujas faces são identificadas por duas translações ortogonais, que são geradores do grupo de isometrias Γ_{T^2} . Como todo o grupo é composto por translações, ele é globalmente homogêneo, e a forma do domínio fundamental não depende da escolha do ponto base. No caso da garrafa de Klein, o grupo Γ_K é gerado por uma translação, e por uma isometria 'tipo parafuso', que consiste da composição de uma rotação (de π , no caso) em torno de um eixo e uma translação ao longo do mesmo. Esta última isometria não comuta com translações perpendiculares ao seu eixo, e portanto a forma do domínio fundamental

⁹Dizemos que α comuta com β se $\alpha = \beta\alpha\beta^{-1}$. Obviamente não podemos definir um operador de comutação $[\alpha, \beta]$, pois tratamos de um grupo, não de uma álgebra.

depende do ponto base. Note ainda que a composição dupla do gerador 'tipo parafuso' é uma translação.

Capítulo 2

Anisotropia estatística na radiação cósmica de fundo

Existe uma teoria segundo a qual se alguém descobrir exatamente o que é e para que serve o Universo, este irá desaparecer imediatamente, e em seu lugar surgirá algo ainda mais bizarro e inexplicável. Outra teoria diz que isso já aconteceu.

Douglas Adams

A minha suspeita é que o Universo não só é mais estranho do que supomos, mas sim mais estranho do que somos capazes de supor.

J.B.S. Haldane

2.1 Introdução

As análises mais recentes da radiação cósmica de fundo (RCF) têm sugerido que o Universo observável não é estatisticamente isotrópico em grandes escalas. Se for comprovada, tal anisotropia terá consequências profundas para o nosso entendimento do cosmo, já que o modelo cosmológico padrão e os pressupostos observacionais usuais baseiam-se na premissa de que o Universo é isotrópico e homogêneo em grandes escalas. Não há

porém consenso sobre a significância dos indícios de tal anisotropia, e as hipóteses para sua origem ainda são meramente especulativas. Assim, é importante procurar por indicadores que possam determinar quantitativamente a anisotropia, indicar se sua origem é cosmológica ou alguma fonte local, e, se possível, dar pistas sobre a sua origem. Com isso em mente, propomos neste capítulo um indicador direcional construído a partir de mapas da RCF, e baseado em um coeficiente que mede o desvio da isotropia estatística em histogramas de separação de pares. A principal vantagem deste indicador é que, a partir de um mapa de céu inteiro das anisotropias locais, ele produz um mapa derivado anisotropia global. Quando aplicado às medições mais recentes da radiação cósmica de fundo, nossa abordagem indica a existência de uma anisotropia estatística, concentrada em uma direção específica. A comparação com mapas gerados pelo método de Monte Carlo a partir da hipótese de isotropia estatística mostra que a anisotropia global detectada é estatisticamente significativa em grandes escalas angulares.

Como veremos, nossos resultados são robustos em relação à escolha dos parâmetros do método, desde que o ruído estatístico seja bem controlado. A escolha do algoritmo para filtragem do foreground¹ usado para gerar os mapas de RCF tampouco influi de forma significativa nos resultados finais.

Existe uma profusão de métodos propostos na literatura para detectar anisotropias globais. Alguns não acham nenhum desvio significativo, outros encontram diversos eixos e direções preferenciais. Dada esta Babel de resultados, é difícil afirmar se a anisotropia que encontramos coincide ou não com outros resultados na literatura, embora em alguns casos a evidência seja sugestiva. Porém, o nosso método é ao mesmo tempo, e de forma única, quantitativo e direcional. Isto nos permite ir além da mera constatação da existência da anisotropia global ou da presença de direções preferenciais, para mapear a anisotropia global (ou pelo menos uma de suas facetas) no céu. Inicialmente o método foi aplicado a mapas de céu inteiro [16], e depois estendido para mapas em que as regiões de alta contaminação foram removidas por meio do uso de máscaras [114].

2.2 O sucesso da cosmologia, e algumas pequenas discrepâncias

Apesar do grande sucesso do modelo padrão da cosmologia em prever ou explicar os mapas de RCF, existem ainda algumas possíveis anomalias em grandes escalas angulares que não são explicadas por um modelo de aceitação geral (para uma revisão geral

¹Usaremos o termo original em inglês por não haver correspondente de uso corrente em português.

veja [35, 36]). Tais anomalias incluem um valor excepcionalmente baixo dos momentos de quadrupolo e octopolo ([157], [156]), o alinhamento dos eixos de quadrupolo e octopolo ([164], [44], [181], [22], [185] e [1]), evidências sugerindo uma assimetria entre o norte e o sul galácticos ([73], [53] e [19]), indicações da existência de um eixo preferencial de simetria na direção $(b = 30^\circ, l = 250^\circ)^2$, e uma direção de máxima assimetria em $(b = 100^\circ, l = 237^\circ)$ ([26], [34], [53], [73, 74], [186], [89, 90], [149], [136], [48]). Existem ainda indicações de não-gaussianidade nas flutuações de temperatura da RCF ([34], [87], [125], [30], [175], [52], [91], [32] e [18]). Todos estes casos têm em comum o fato de que estatisticamente a RCF não parece ter exatamente as mesmas propriedades em todas as direções.

A mera constatação da existência de tais anomalias obviamente não resulta em conhecimento de sua(s) origem(s), que pode ser cosmológica, ou o resultado de contaminação de foreground ou erros sistemáticos não corrigidos ([149]). Outras explicações propostas incluem efeitos de lentes gravitacionais (veja [174], e a sua contestação[33]) e mecanismos de quebra da isotropia estatística ([66] e [171]). Existem ainda propostas que contemplam uma origem extra-galáctica do fenômeno ([26], [164], [53], [73] e [89, 90]).

No caso de se confirmar uma origem cosmológica para a anisotropia revelada por estas análises, seria necessário questionar algumas das premissas básicas do modelo padrão da cosmologia. Em particular, será preciso revisar o paradigma inflacionário, que prevê [163] gaussianidade e isotropia estatística nas flutuações de temperatura da RCF.

Por causa disto, foram feitas várias tentativas de verificar a existência de tais anomalias por meio de diversos métodos estatísticos. Hajian e Souradeep, por exemplo, propuseram uma medida quantitativa, mas não direcional, da anisotropia estatística [155], que sugerem ser útil para detectar indícios de uma eventual topologia não trivial para o Universo. Além da vantagem de corroboração mútua, o uso de vários métodos também pode fornecer informação sobre diferentes tipos de anisotropias, ou diferentes facetas do mesmo fenômeno. Também é razoável esperar que a combinação de múltiplos métodos torne possível determinar a origem, ao invés de simplesmente constatar a presença, da(s) anisotropia(s) (c.f. novamente os trabalhos de Hajian e Souradeep, e em particular [154], [71], [70] e [72]).

²Usamos novamente coordenadas galácticas com o pólo norte galáctico em $b = 0^\circ$, o equador definido por $b = 90^\circ$, o centro galáctico em $l = 0^\circ$ e $l \in [0^\circ, 360^\circ)$ crescendo do oeste para o leste.

2.3 Radiação cósmica de fundo e anisotropia estatística

Com o avanço da cosmologia de precisão, nos últimos anos tem ocorrido um enorme aumento da quantidade e qualidade dos dados disponíveis para análise. Em particular, os dados referentes a três anos de observações contínuas da radiação cósmica de fundo coletadas pelo observatório orbital WMAP (Wilkinson Microwave Anisotropy Probe ([11, 10] e [77])) produziram um salto qualitativo no nosso entendimento do Universo primordial. Tais dados indicam com considerável precisão que o Universo é espacialmente plano ou quase plano, com um espectro primordial de perturbações de densidade quase invariante por escala. Estes resultados correspondem às previsões do modelo cosmológico padrão.

Uma premissa fundamental (e muitas vezes aceita implicitamente) da cosmologia é que a nossa posição no cosmos não é excepcional. De forma um pouco mais geral, a cosmologia padrão supõe que todos os pontos do Universo observável são equivalentes, ou seja, que o Universo é homogêneo em grandes escalas, e de fato até onde nos é dado a ver, isto se verifica em escalas $\gtrsim 100 \text{ Mpc}$. Obviamente, se o Universo fosse precisamente homogêneo espacialmente, não haveriam galáxias, nem seres humanos para observá-las.

A RCF é hoje um dos nossos principais instrumentos para estudar as pequenas flutuações da homogeneidade que deram origem às estruturas que observamos hoje. Esta radiação, dita fóssil, é um retrato (quase) instantâneo de quando o Universo havia se expandido o suficiente para que o plasma que o preenchia esfriasse e se combinasse em átomos, permitindo então a passagem da radiação térmica que até então estava em equilíbrio térmico com a matéria. O evento do desacoplamento define uma seção espacial da 4-variedade riemanianna utilizada para modelar o Universo, chamada hiper-superfície de último espalhamento (HSUE); a interseção da HSUE com o nosso cone de luz passado (ou seja, a parte que incide sobre nossa posição a cada instante) é a superfície de último espalhamento, ou SUE.

Como o princípio da homogeneidade implica que o desacoplamento ocorreu na mesma época por todo o Universo, e a radiação então se propagou livremente até ser detectada por nossos instrumentos, a SUE toma a forma de uma esfera (para um observador comóvel com o fluxo de Hubble! Veja o Cap. 5), cujo raio medido em anos-luz corresponde à idade do Universo menos o tempo necessário para que a recombinação ocorresse.

Quando foi emitida, portanto, a RCF era térmica (i.e., seguia, e ainda segue, uma distribuição Planckiana), sendo definida portanto pela temperatura da matéria no local de emissão. Assim, as flutuações que constituem os desvios locais da homogeneidade global presentes na SUE serão manifestas na RCF como flutuações de temperatura. Atualmente, portanto, esta radiação (que permaneceu quase inalterada a menos de um considerável

desvio para o vermelho devido à expansão do Universo), contém informações diretas a respeito de flutuações de densidade da matéria em pontos diferentes (as inomogeniedades), através das concomitantes flutuações, ou anisotropias, de temperatura da radiação vinda de direções diferentes.

A observação da RCF tem sido extremamente proveitosa para a ciência. É possível medir a temperatura da radiação com extrema precisão; e como as flutuações são muito pequenas, da ordem de 10^{-5} , em relação à temperatura média, é possível construir modelos usando física bem conhecida que detalham a evolução das flutuações primordiais lineares até a formação da estrutura em larga escala do Universo atual. O estudo das anisotropias da RCF permite testar detalhadamente tais modelos e suas premissas.

Quase todos os estudos da RCF têm implícitos a premissa de que o Universo é homogêneo em grandes escalas, da ordem de ou maiores que 100 Mpc , e aproximadamente isotrópico. Mas será possível utilizar a própria RCF para testar estas hipóteses?

Sabemos que o Universo não é nem exatamente homogêneo, nem exatamente isotrópico. Para verificar que o Universo é aproximadamente homogêneo é necessário, em princípio, somente verificar que em uma escala suficientemente grande os desvios em relação à homogeneidade são insignificantes. Qual seria o conceito correspondente para 'aproximadamente isotrópico'? Na esfera celeste a escala é fixa, e não existe portanto a noção de 'isotrópico em uma escala suficientemente grande'. O que procuramos é uma definição de isotropia estatística. Posto da forma mais simples possível, um mapa é estatisticamente isotrópico quando todas as direções são estatisticamente equivalentes.

É preciso portanto explicitar o conceito de 'estatisticamente isotrópico', para que seja possível verificar sua presença (ou ausência) na RCF. Para tanto, estudamos brevemente alguns detalhes da estatística das suas flutuações de temperatura. (para um tratamento lúcido do assunto veja Souradeep & Hajian, em e.g. [155, 154])

Em princípio, um mapa da RCF define uma função escalar $T(\hat{n})$ em $\mathbb{S}^2$, onde T é a temperatura da radiação incidente da direção $\hat{n}$. As flutuações da RCF são obviamente definidas por $\Delta T(\hat{n}) = T(\hat{n}) - \langle T \rangle_{\mathbb{S}^2}$, onde a média é tomada por toda $\mathbb{S}^2$ (e indicada por $\langle \rangle_{\mathbb{S}^2}$), e tipicamente $\langle T \rangle \gg \Delta T(\hat{n})$. Podemos considerar $\Delta T(\hat{n})$ como uma função aleatória suave definida na esfera, cuja distribuição de probabilidade é dada por um funcional $P[\Delta T(\hat{n})]$ (veja e.g. [122]), que é totalmente definido por suas funções de correlação (FDC), i.e.,

$$\begin{aligned}
\langle \Delta T(\hat{n}_1) \rangle &= \int \Delta T(\hat{n}_1) P[\Delta T(\hat{n})] DT & (\text{Média}) \\
\langle \Delta T(\hat{n}_1) \Delta T(\hat{n}_2) \rangle &= \int \Delta T(\hat{n}_1) \Delta T(\hat{n}_2) P DT & (\text{Covariância}) \\
&\vdots \\
\langle \Delta T(\hat{n}_1) \Delta T(\hat{n}_2) \dots \Delta T(\hat{n}_N) \rangle &= \int \Delta T(\hat{n}_1) \Delta T(\hat{n}_2) \dots \Delta T(\hat{n}_N) P DT & (\text{FDC de N pontos})
\end{aligned} \tag{2.1}$$

onde $DT = \prod_{\hat{n}' \in \mathbb{S}^2} dT(\hat{n}')$.

A definição formal de isotropia estatística se torna então natural. Um mapa é estatisticamente isotrópico quando é a realização de uma distribuição cujas funções de correlação são invariantes por rotações de $\mathbb{S}^2$.

Existem dois impedimentos ao estudo da RCF usando simplesmente a definição acima. O primeiro, ao qual voltaremos logo adiante, deriva do fato de só existir um mapa da RCF passível de análise, a partir do qual precisamos deduzir todas as infinitas funções de correlação. O segundo impedimento é uma consequência direta deste fato: não sabemos, em princípio, qual a distribuição $P[\Delta T(\hat{n})]$, o que seria necessário antes de realizar qualquer teste direto da isotropia estatística.

Este último problema pode ser contornado, se a distribuição puder ser deduzida a partir dos modelos para as flutuações primordiais. E de fato, acredita-se que tais flutuações tenham origem em flutuações quânticas amplificadas pelo processo inflacionário [94], que dariam origem a flutuações gaussianas, ou seja [122]

$$P[\Delta T(\hat{n})] = \frac{1}{\mathcal{N}} \exp \left[-\frac{1}{2} \int C^{-1}(\hat{n}, \hat{n}') \Delta T(\hat{n}) \Delta T(\hat{n}') d\Omega d\Omega' \right], \tag{2.2}$$

onde $\mathcal{N}$ é o fator de normalização, e

$$\langle \Delta T(\hat{n}) \Delta T(\hat{n}') \rangle = C(\hat{n}, \hat{n}') . \tag{2.3}$$

Podemos ainda expressar as flutuações de temperatura em termos dos harmônicos esféricos usuais,

$$\Delta T(\hat{n}) = \sum_{l=\ell}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_l^m, \tag{2.4}$$

$$C_\ell = \frac{1}{2\ell + 1} \sum_m |a_{\ell m}|^2. \tag{2.5}$$

Em uma distribuição gaussiana de média zero, todas as demais funções de correlação são determinadas a partir da função de correlação de dois pontos, $C(\hat{n}, \hat{n}')$ (também chamada de covariância). Além disso, neste caso é fácil ver que no caso de isotropia estatística, $C(\hat{n}, \hat{n}') = C(\hat{n} \cdot \hat{n}')$; ou seja, a covariância só depende da separação angular entre os pontos. Lembrando ainda que o espectro de potência angular denotado pelos C_ℓ é invariante por rotações (o que não é o caso para os $a_{\ell m}$!), é possível mostrar que isotropia estatística ocorre se e somente se a covariância pode ser expressa em termos dos C_ℓ da seguinte maneira

$$C(\hat{n} \cdot \hat{n}') = \frac{1}{4\pi} \sum_{\ell} (2\ell + 1) C_\ell P_\ell(\hat{n} \cdot \hat{n}') , \quad (2.6)$$

em que P_ℓ denotam os polinômios de Legendre. Como distribuições gaussianas são totalmente determinadas por sua covariância, é possível obter a partir de (2.6) toda a informação contida na distribuição (2.2).

De posse da fórmula (2.6), é aparentemente fácil verificar se um dado mapa da RCF é estatisticamente isotrópico ou não. Bastaria estimar a função de correlação de dois pontos para a RCF e comparar as duas funções. Mas na realidade esbarramos na primeira objeção levantada acima. Como só há um céu de microondas para ser observado, só podemos obter uma realização da covariância, e a diferença entre ela e (2.6) permanece resolutamente dentro dos limites da variância cósmica. Para a RCF, este método direto simplesmente não é sensível o suficiente.

Assim, devido à variância cósmica, não é possível determinar se a RCF apresenta isotropia estatística diretamente a partir do espectro de potência, exceto em casos extremos. Outros métodos mais elaborados se tornam portanto necessários.

2.4 O σ -mapa

Nos mapas de céu inteiro usados no estudo da RCF a esfera celeste é discretizada em um certo número de pixels de mesma área, e a cada pixel é associado um valor de temperatura (para uma discussão completa sobre métodos de pixelização, veja [123]). Partindo de um destes mapas, inicialmente ordenamos os pixels por sua temperatura, e dividimos o mapa em um conjunto de submapas, que consistem cada um em um número constante de pixels de temperaturas consecutivas. Cada submapa consiste então no conjunto de pixels, espalhados pela esfera celeste, que têm temperaturas situada em uma dada faixa. Como veremos adiante, o número exato de submapas usados não influi significativamente nos resultados obtidos, desde que certas restrições bem definidas sejam respeitadas.

A motivação para tal subdivisão vem do fato de que em um mapa estatisticamente isotrópico, pixels de todas as faixas de temperatura deverão estar distribuídos de forma homogênea no céu. Da mesma forma, se houver alguma anisotropia, pixels 'quentes' e 'frios' não devem se distribuir da mesma forma (em mapas com alto grau de anisotropia, como por exemplo um dipolo puro, eles se concentram em regiões diferentes do céu). Como não sabemos a priori quais tipos de anisotropia esperar, este procedimento tem a vantagem de especificar claramente o resultado esperado no caso estatisticamente isotrópico.

2.4.1 O histograma de separação angular de pares

Para quantificar a anisotropia em uma região da esfera celeste, definimos os chamados histogramas de separação angular de pares (HSAP, ou PASH em inglês), subdividindo o segmento $\alpha \in (0, 180^\circ]$ em N_{bins} intervalos $J_i = (\alpha_i - \frac{\delta\alpha}{2}, \alpha_i + \frac{\delta\alpha}{2}]$. Contamos então o número de pares de pixels em cada submapa cuja separação se situa em cada um destes intervalos. A partir da seqüência de valores assim obtida o HASP é então definido por

$$\Phi(\alpha_i) = \frac{1}{N\delta\alpha} \sum_{\alpha \in J_i} \eta(\alpha), \quad (2.7)$$

sendo que N é o número de pares de pixels no submapa, $\eta(\alpha)$ é o número de pares de pixels com separação α , e $\delta\alpha = \pi/N_{\text{bins}}$ é o comprimento dos intervalos. A normalização garante que a condição $\sum_{i=1}^{N_{\text{bins}}} \Phi(\alpha_i) \delta\alpha = 1$ seja válida. Calculamos então o valor médio em cada um dos subintervalos dos HSAPs de todos os submapas, para obter o HSAP médio (HSAPM, ou MPASH em inglês), $\langle \Phi(\alpha_i) \rangle$.

Para obter uma medida da anisotropia para um mapa discreto (i.e., pixelizado), comparamos o HSAPM calculado da forma descrita acima com o histograma médio esperado em mapas estatisticamente isotrópicos. A definição do HSAPM, acima, se aplica a qualquer região da esfera celeste, inclusive à esfera inteira, e obviamente a função Φ_{exp} depende da forma da região. Mas para os nossos fins neste trabalho, só precisamos considerar calotas esféricas, i.e., regiões definidas como o interior de uma circunfeência definida na superfície de $\mathbb{S}^2$ (a esfera completa, a menos de um ponto de medida zero, é um caso degenerado de calota). Para calotas de abertura ρ , existe uma fórmula obtida por Teixeira ([166], veja também [14] para uma derivação numérica alternativa) que nos dá a distribuição de probabilidade de pares de pontos com separação α

$$\mathcal{P}_{\text{exp}}(\alpha; \rho) = \frac{\sin \alpha}{\pi (1 - \cos \rho)^2} \left\{ \sin^{-1} \left[\frac{(\cos \rho + \cos \alpha) \sqrt{\cos \alpha - \cos(2\rho)}}{\sin \alpha_0 (1 + \cos \rho) \sqrt{1 + \cos \alpha}} \right] + (1 - 2 \cos \rho) \cos^{-1} \left[\cot \rho \tan\left(\frac{\alpha}{2}\right) \right] \right\}. \quad (2.8)$$

onde $\alpha \in (0, 2\alpha_0]$, $\alpha_0 \leq \pi$. No caso da esfera inteira, a fórmula acima se reduz a $\mathcal{P}_{\text{exp}}(\alpha; \pi) = \frac{1}{2} \sin \alpha$.

Definimos assim o histograma de separação de pares esperado (HSAPE) em calotas, $\Phi_{\text{exp}}(\alpha_i)$, a partir dos valores assumidos pela distribuição $\mathcal{P}_{\text{exp}}$ no centro dos intervalos J_i

$$\Phi(\alpha_i) = \frac{1}{\delta\alpha} \mathcal{P}_{\text{exp}}(\alpha_i; \rho); \quad (2.9)$$

deixamos implícita a dependência da abertura ρ .

Podemos definir portanto como um indicador de anisotropia o histograma da diferença entre o HSAPM (2.7) de uma calota de um mapa pixelizado observacional, e o HSAPE (2.9) para uma calota de mesmo raio

$$\Upsilon(\alpha_i) \equiv \langle \Phi_{\text{obs}}(\alpha_i) \rangle - \Phi_{\text{exp}}(\alpha_i). \quad (2.10)$$

Definido desta maneira, é evidente que o histograma de diferença Υ contém informação sobre a amplitude e escala angular da anisotropia na região da esfera em que é calculado, que daqui para a frente sempre vamos supor ser uma calota de abertura ρ e centro em $\hat{c}$. O histograma Υ obtido para esta calota mede assim a anisotropia em uma vizinhança de $\hat{c}$. De forma mais direta, mas um pouco menos precisa, podemos dizer que Υ é uma medida da anisotropia na direção $\hat{c}$.³

Para ilustrar as afirmativas acima, mostramos na Fig. 2.1 os histogramas de diferença correspondentes a duas calotas antipodais (com abertura $\rho = 30^\circ$), com centros em $(b = 65^\circ, l = 56^\circ)$ e $(b = 115^\circ, l = 236^\circ)$, calculados para o mapa ILC [77] da RCF. Esta direção foi escolhida explicitamente para mostrar o contraste em regiões de alta e baixa anisotropia. Vemos claramente que a primeira calota se desvia da isotropia em várias escalas angulares de forma estatisticamente significativa (i.e., além de dois desvios-padrão da média de calotas estatisticamente isotrópicas). A segunda calota, por outro lado, não apresenta desvios significativos.

³Formalmente, é possível provar que quando tomamos $\rho = \pi$, o indicador Υ corresponde formalmente à função de correlação de dois pontos para mapas estatisticamente isotrópicos ([15]), a menos de um fator de normalização.

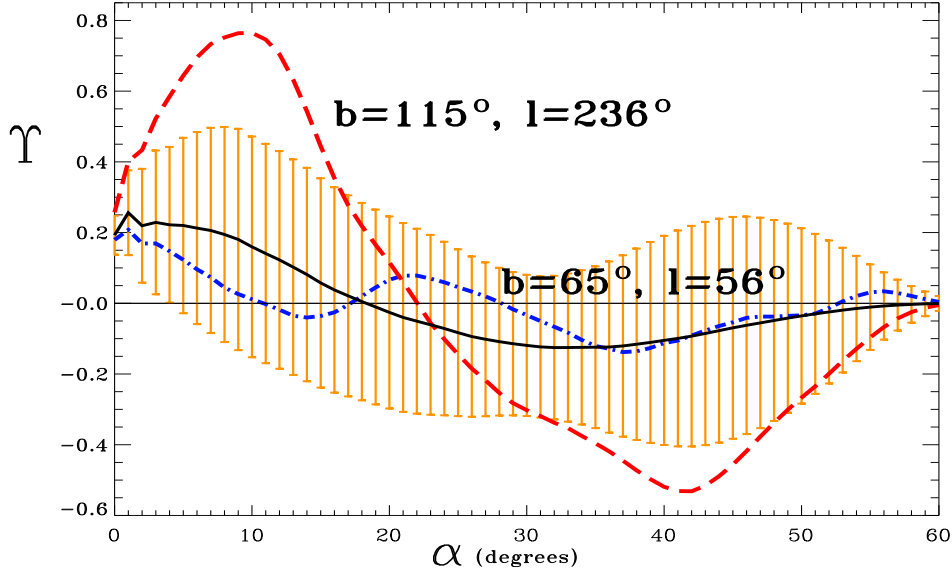


Figura 2.1: Exemplos do histograma de diferença $\Upsilon(\alpha)$ em duas calotas do mapa LILC-ILC diametralmente opostas, centradas em $(b = 65^\circ, l = 56^\circ)$ e $(b = 115^\circ, l = 236^\circ)$, com $\gamma_0 = 30^\circ$. O valor médio de Υ para o mapa em questão é indicado pela linha sólida, e o sombreado corresponde ao limite de dois desvios-padrão. Fica claro que a anisotropia estatística é significativa apenas em uma das calotas.

2.4.2 O indicador σ

O estudo de calotas individuais poderia indicar a presença de anisotropias em grandes escalas. Mas para uma análise sistemática da dependência direcional da anisotropia, é preciso quantificar como a anisotropia varia ao longo de toda a esfera. A maneira mais natural de definir tal medida é construindo uma função escalar (e não um histograma) definida em toda a esfera celeste, que assume valores em cada ponto correspondentes à escala da anisotropia medida em uma vizinhança daquela direção. Continuando o procedimento descrito acima, este indicador para cada ponto será obtido a partir do histograma de diferença Υ de uma calota centrada na direção correspondente.

Para definir este indicador de forma concreta, sejam $\Omega_j \equiv \Omega(\theta_j, \phi_j; \rho) \in \mathbb{S}^2$ um conjunto de calotas esféricas com raio ρ , e cujos centros (θ_j, ϕ_j) , $j = 1, \dots, N_{\text{cals}}$ correspondem a uma pixelização do céu, e onde a união das N_{cals} calotas cobre toda a esfera celeste $\mathbb{S}^2$. Definimos a função escalar $\sigma : \Omega_j \mapsto \mathbb{R}^+$, que associa à j -ésima calota um número real positivo σ_j dado pela variância de Υ (cujo valor médio é zero por construção), calculado na calota, i.e.:

$$\sigma_j^2 \equiv \frac{1}{N_{\text{bins}}} \sum_{i=1}^{N_{\text{bins}}} \Upsilon_j^2(\alpha_i), \quad (2.11)$$

onde $\Upsilon_j \equiv \Upsilon(\theta_j, \phi_j)$ é o histograma Υ calculado para a calota com centro em (θ_j, ϕ_j) .

Coletando os valores de σ para cada calota, a função $\sigma = \sigma(\theta, \phi)$, tendo como domínio a totalidade da esfera celeste,⁴ mede desvios da isotropia como função da direção (θ, ϕ) . Assim, podemos dizer que $\sigma(\theta, \phi)$ é uma medida direcional da anisotropia sobre a esfera celeste.

Note que, como $\sigma(\theta, \phi)$ é uma função escalar definida em $\mathbb{S}^2$, ela pode ser também entendida como um mapa, que denominamos σ -mapa. A forma mais natural de quantificar os detalhes da informação sobre anisotropias nela contidos é, assim como no caso das flutuações de temperatura, efetuar uma expansão em harmônicos esféricos,

$$\sigma(\theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} b_{\ell m} Y_{\ell m}(\theta, \phi) . \quad (2.12)$$

e calcular o espectro de potência correspondente

$$D_{\ell} = \frac{1}{2\ell + 1} \sum_m |b_{\ell m}|^2 . \quad (2.13)$$

Se alguma anisotropia em grandes escalas estiver presente no mapa de flutuações de temperatura original, ela deve afetar também o σ -mapa em escalas angulares comparáveis, e portanto este efeito deve ser mais evidente nos multipolos mais baixos.

Recapitulando, o processo para gerar um σ -mapa a partir de um mapa pixelizado de flutuações de temperatura consiste nas seguintes etapas:

- (i) Gerar um conjunto de calotas distribuídas homoganeamente sobre a esfera celeste;
- (ii) Subdividir cada calota em N_{his} submapas, cada um contendo os pixels situados em uma faixa de temperatura escolhida de modo que todos os submapas tenham um número igual de pixels;
- (iii) Calcular o HSAP Φ_{obs} correspondente a cada submapa;
- (iv) Calcular a média de todos os Φ_{obs} em cada calota para obter o $\langle \Phi_{\text{obs}} \rangle$;
- (v) Subtrair $\Phi_{\text{exp}} - \langle \Phi_{\text{obs}} \rangle$ para obter o histograma da diferença, Υ ;
- (vi) Calcular o valor de σ para cada calota, a partir do histograma Υ ;
- (vii) Definir o σ -mapa como uma função escalar $\sigma(\theta, \phi)$ em $\mathbb{S}^2$, definida no centro de cada calota;
- (viii) Calcular os multipolos D_{ℓ} para o σ -mapa.

⁴Ou, no caso prático discretizado, o conjunto dos centros dos pixels que cobrem toda a esfera celeste.

2.5 Aplicação aos dados da RCF

Lançado em 2001, o observatório orbital de microondas WMAP produziu mapas de alta resolução das flutuações de temperatura da RCF. Os mais recentes dados disponíveis são referentes a três anos de observação continuada, e foram publicados em março de 2006 [10]. O WMAP observa o céu a partir do ponto de Lagrange L2, situado a 1.5 milhões de quilômetros da Terra, na direção diametralmente oposta ao Sol. Desta maneira, é possível observar continuamente a RCF sem interferências ou obstruções do Sol, Terra e Lua. À medida em que nosso planeta se move em sua órbita, o WMAP observa a totalidade do céu a cada seis meses.

O antecessor imediato do WMAP, o satélite COBE, foi um dos precursores da chamada cosmologia de precisão [153]. O WMAP [188] ambicionava repetir o mapeamento do céu de microondas com maior resolução, e uma grande ênfase no controle sobre erros sistemáticos. Suas metas observacionais, plenamente atingidas, incluíam uma resolução angular de pelo menos 0.3° , e sensibilidade de $20\mu K$ por pixel. A contribuição de artefatos observacionais gerados por erros sistemáticos deveria ser menor do que $5\mu K$ por pixel.

2.5.1 A escolha de mapas e a contaminação de foreground

Os dados brutos do WMAP consistem em observações ininterruptas em cinco faixas de frequência, (centradas em 22, 33, 40, 60 e 93 GHz, convencionalmente chamadas de bandas K, Ka, Q, V, and W). Estes dados são combinados para produzir mapas de RCF, e normalmente são tratados para remover o componente de dipolo devido ao movimento próprio do satélite em relação ao referencial comóvel local, e contaminações de foreground (que se concentra no plano galáctico, e em particular na direção do centro da Via Láctea em Sargitário). Existem vários métodos de limpeza em uso, que variam em complexidade, mas são em geral bastante não-triviais, para assegurar que a contaminação seja removida sem que artefatos do tratamento de dados sejam introduzidos. Os mapas pixelizados resultantes são a base a partir da qual geramos os σ -mapas. Utilizamos aqui os seguintes mapas referentes ao dados completos de três anos: O mapa dito COADDED [77] (que consiste simplesmente em uma combinação linear dos mapas obtidos a partir das faixas de frequência usadas na observação, que busca minimizar mas não remover a contaminação galáctica), e os mapas limpos ILC [77], OT [43] e PPG [126].

Para fins de comparação, e para estudar a robustez do método, estudamos ainda mapas referentes aos dados incompletos de somente um ano de observações, o mapa Lagrange-ILC (LILC) [51], TOH [164], além do mapa COADDED [78].

Normalmente os mapas tratados da RCF são pixelizados por meio do algoritmo HE-

ALPix (Hierarchical Equal Area isoLatitude Pixelization; para uma explicação detalhada veja [67]), que define pixels de áreas iguais e formas aproximadamente retangulares, distribuídos de forma razoavelmente homogênea na esfera celeste. A resolução dos mapas HEALPix pode ser ajustada pela fusão ou subdivisão dos pixels, e é especificada pelo parâmetro N_{side} . O número total de pixels é dado por $12 \times N_{\text{side}}^2$. Para os mapas descritos acima usamos sempre a resolução $N_{\text{side}} = 128$, que corresponde a uma partição da esfera celeste em 196608 pixels.

Como veremos a seguir, os resultados obtidos em todos os mapas, e em especial nos mapas limpos, são bastante similares entre si. Assim, apresentamos uma análise mais detalhada somente do mapa de três anos ILC, e uma análise comparativa com os demais mapas.

2.5.2 Calculando os σ -mapas

Para calcular⁵ os σ -mapas correspondentes aos mapas descritos acima, utilizamos $N_{\text{cals}} = 12288$ calotas com 30° de abertura, com os centros distribuídos de acordo com o parâmetro de HEALPix $N_{\text{side}} = 32$. Cada calota foi subdividida em $N_{\text{his}} = 64$ submapas, com o mesmo número de pixels, cada um correspondendo a uma faixa de temperaturas. O número de subdivisões nos histogramas de separação de pares é sempre $N_{\text{bins}} = 180$. Estes valores foram escolhidos para maximizar a resolução dos σ -mapas dados os recursos computacionais disponíveis. Mas como veremos a seguir, a mudança nos valores dos parâmetros, desde que respeitados certos limites bem entendidos, não altera de forma significativa os resultados finais.

A Figura 2.2 mostra, em coordenadas galácticas, a projeção Mollweide do σ -mapa obtido a partir do mapa de flutuações de temperatura da RCF denominado ILF. Este σ -mapa mostra claramente uma região de alto valor de σ entre o centro galáctico e o quadrante sudoeste (direito e inferior nas figuras). Grosso modo, tal região sugere uma distribuição essencialmente dipolar de anisotropia. Em particular, existe uma mancha com valores muito altos de σ concentrados próximo a um máximo bem definido em $(b = 108^\circ, l = 222^\circ)$. Este máximo fica a 16.6° de separação dos eixos preferenciais encontrados recentemente por Eriksen *et al.* [53] e Land & Magueijo [89]. Observa-se ainda uma mancha anular menos proeminente, muito próxima do centro da galáxia.

Para ir além da descrição qualitativa do σ -mapa, precisamos quantificar a informação sobre a direção e intensidade da anisotropia observada. Assim como usualmente é feito

⁵As rotinas usadas para calcular os σ -mapas e seus espectros angulares de potência foram desenvolvidas na linguagem IDL, e estão disponíveis mediante pedido.

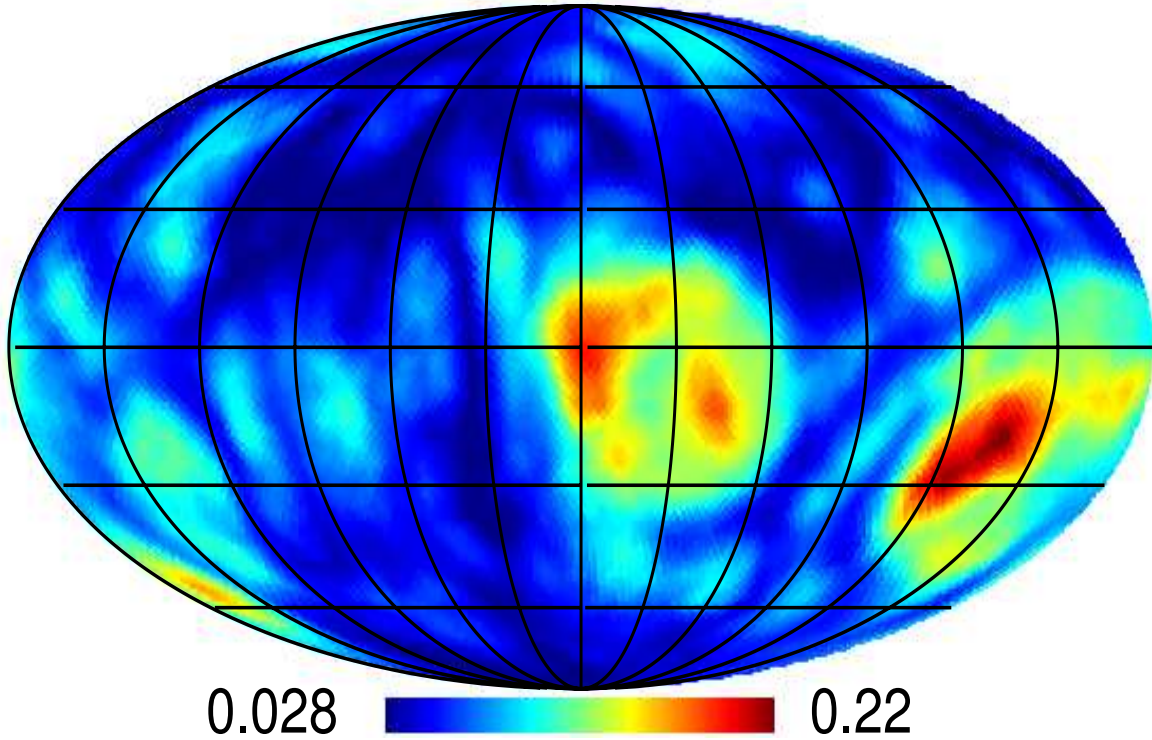


Figura 2.2: O σ -mapa de céu inteiro obtido a partir do mapa ILC, calculado para 12288 calotas com 30° de raio. Nesta imagem, como nas subseqüentes, o pólo norte galáctico fica no topo, e a direção leste-oeste segue da esquerda para a direita.

com os mapas de temperatura, calculamos o espectro de potência do σ -mapa obtido a partir do mapa ILC [77] da RCF. Note, porém, que como obviamente os σ -mapas não são estatisticamente isotrópicos, não é verdadeiro afirmar que toda informação estatisticamente relevante sobre os componentes $b_{\ell m}$ está contida nos valores de D_ℓ . O espectro de potência angular, D_ℓ , é representado por pontos sólidos na Figura 2.3. Note os altos valores do dipolo, quadrupolo e octopolo, que deixam claro que a anisotropia observada na Figura 2.2 se dá primariamente em grandes escalas angulares ($\ell \leq 4$).

2.5.3 O espectro angular de potência e a relevância estatística dos resultados

Continuando nossa análise, estudamos o significado estatístico dos valores observados dos σ -multipolos derivados do mapa ILC. Foram gerados 1000 mapas simulados da RCF estatisticamente isotrópicos. Cada mapa simulado da RCF gerado pelo método de Monte Carlo é uma realização estocástica do melhor ajuste para o espectro de potência angular do WMAP (assumindo um modelo Λ CDM [156]), obtido a partir de componentes randômicos $a_{\ell m}$ gerados dentro dos limites da variância cósmica.

Verificamos numericamente que a distribuição dos valores dos componentes $b_{\ell m}$ nos σ -mapas obtidos a partir dos mapas de RCF simulados seguem distribuições aproximadamente gaussianas. Como vimos na Seção 1.3.4, isto implica que cada componente do espectro angular de potência associado, D_ℓ , deve seguir uma distribuição de χ^2 com $(2\ell + 1)$ graus de liberdade. Esta é exatamente a definição da variância cósmica utilizada em mapas de RCF (quando se parte da premissa que estes são estatisticamente isotrópicos). Assim, podemos afirmar sem abuso de linguagem que também existe uma variância cósmica induzida nos σ -mapas.

Os valores de ℓ usados nesta comparação vão de $\ell = 1$ até $\ell = 10$, o que permite resolução suficiente para investigar as correlações angulares em grandes escalas correspondentes a separações da ordem de $\gtrsim \pi/\ell \simeq 18^\circ$ no σ -mapa. Note que os valores de D_ℓ caem rapidamente para $\ell \geq 4$, e convergem para os valores médios nos mapas simulados estatisticamente isotrópicos. Na Figura 2.3 o espectro correspondente à média dos espectros de 1000 σ -mapas obtidos a partir de mapas de temperatura estatisticamente isotrópicos simulados por Monte Carlo é indicado pela linha contínua. As regiões hachuradas correspondem à variância cósmica.

Note o valor relativamente alto do dipolo, consistente com o que pode ser constatado visualmente na Figura 2.2. De fato, podemos integrar a distribuição χ^2 (com 3 graus de liberdade) entre 0 e o valor medido para D_1 no σ -mapa obtido do mapa ILC, para calcular a significância estatística do dipolo medido. Verificamos que este valor é estatisticamente significativo com um nível de confiança de 97%. Da mesma maneira, os altos valores do quadrupolo e octopolo e hexadecopolo são significativos com confiança de respectivamente 99%, 99,999% e 99%, respectivamente. A probabilidade de todos os quatro primeiros multipolos terem valores tão altos simultaneamente em um mapa estatisticamente isotrópico pode ser obtida da mesma maneira, e é menor que uma parte em 10^6 . Todos os multipolos subsequentes, exceto (por pouco) D_6 , se situam dentro dos limites da variância cósmica induzida para σ -mapas.

Para verificar a robustez dos resultados apresentados acima, também calculamos o espectro de potência do σ -mapa obtido a partir do mapa PPG [126]. Verificamos que os primeiros três multipolos também apresentam valores significativamente acima do que seria de se esperar no caso estatisticamente isotrópico, todos com significância estatística acima dos 97%.

Dando continuidade à análise da anisotropia encontrada no σ -mapa, analisamos em maiores detalhes a forma dos multipolos anômalos. Na Figura 2.4 estão os componentes de dipolo, quadrupolo e octopolo, além σ -mapa resultante da remoção destes três componentes. O dipolo tem o máximo apontando na direção ($b = 130^\circ, l = 259^\circ$), que por si só

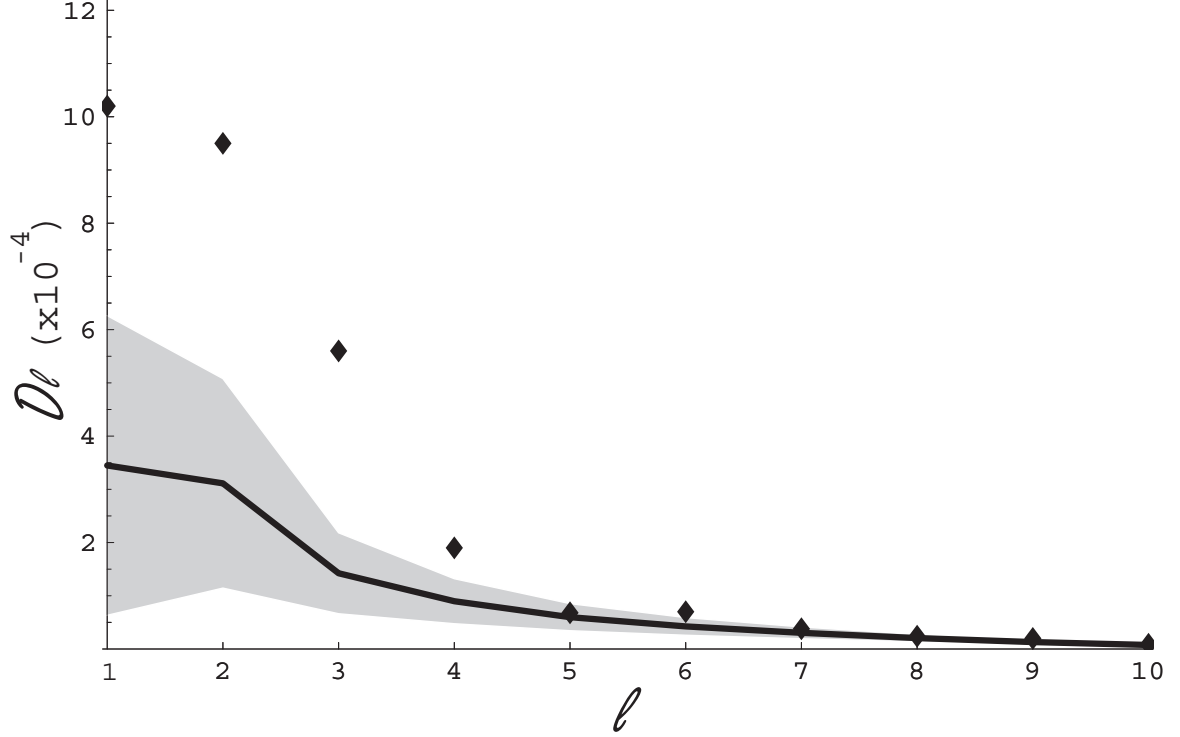


Figura 2.3: O espectro de potência angular do mapa ILC, para $\ell = 1, \dots, 10$ (pontos cheios). Note os altos valores para os três multipolos iniciais. Para comparação, também são mostrados (linha sólida) os valores esperados, obtidos tomando a média entre 1000 mapas de RCF estatisticamente isotrópicos gerados por Monte-Carlo. Os limites da variância cósmica derivada para os σ -maps correspondem à área hachurada.

não é muito próximo dos eixos de máxima assimetria encontrados por Eriksen *et al.* [53] e por Land & Magueijo [89]. Porém, a soma dos três primeiros multipolos resulta em um mapa que já contém as principais características do σ -mapa completo, com uma região anelar de alta anisotropia próximo ao centro galáctico ($b = 97^\circ, l = 342^\circ$), cuja origem provavelmente é o foreground galáctico não removido, e uma mancha com um máximo intenso e bem definido, na direção ($b = 112^\circ, l = 217^\circ$), que está mais próxima aos eixos encontrados por Eriksen *et al.* e Land & Magueijo.

Além do mapa de RCF, vários parâmetros precisam ser especificados para calcular o indicador de anisotropia σ . É importante portanto determinar em que medida a escolha de seus valores pode afetar nossas conclusões. Caso os resultados não sejam robustos, e variem significativamente para escolhas razoáveis dos parâmetros, provavelmente pouco poderia ser afirmado com segurança baseado nos σ -mapas. É fácil porém testar tal robustez, gerando vários σ -mapas a partir do mesmo mapa de RCF, usando diferentes valores de parâmetros.

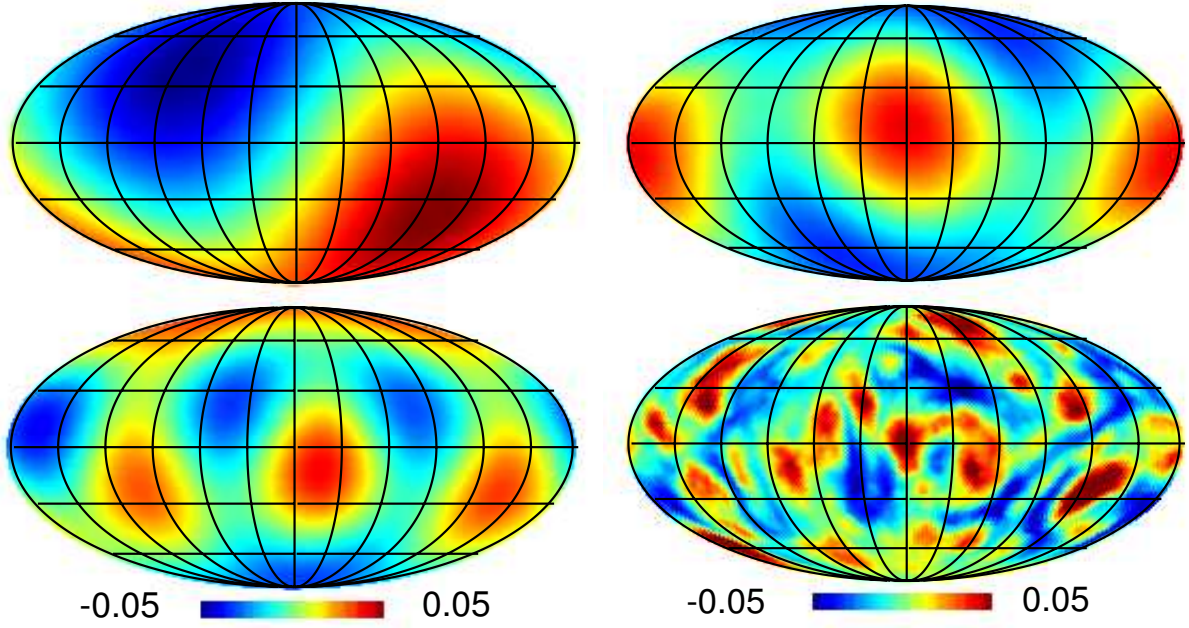


Figura 2.4: Da esquerda para a direita e de cima para baixo estão o dipolo, quadrupolo e octopolo do mapa ILC, assim como o mapa resultante da remoção destes três componentes.

2.5.4 Escolha de parâmetros e robustez

Todos os mapas de RCF considerados têm a resolução do parâmetro HEALpix $N_{\text{side}} = 128$, ou 196608 pixels. Valores maiores de N_{side} para os mapas de RCF não alteram de maneira mensurável os resultados aqui apresentados, e aumentam significativamente o tempo computacional despendido. De fato, os σ -mapas correspondentes foram gerados com $N_{\text{side}} = 32$, que permite visualizar pequenos detalhes das variações de anisotropia. Visualmente, resoluções maiores não produzem mudanças observáveis nos mapas gerados. Quantitativamente, estamos interessados principalmente nas suas características em grandes escalas angulares. Resoluções mais baixas obviamente ocasionam o borramento de detalhes finos, mas para valores maiores que $N_{\text{side}} = 8$ os multipolos de ordem mais baixa não se alteram significativamente. Este fato nos permite utilizar uma resolução mais baixa $N_{\text{side}} = 8$ (o que corresponde a 768 calotas) para gerar os mapas de Monte Carlo, já que eles só são usados para obter valores de D_ℓ esperado para $\ell \leq 10$; computacionalmente valores mais altos para N seriam impraticáveis devido ao considerável tempo de computação que seria necessário para gerar cada mapa.

Dado que $N_{\text{side}} = 32$, para um número de submapas de temperatura N_{his} entre 2 a 64, e um número de subdivisões dos histogramas entre 180 e 400, não há nenhuma variação mensurável nos σ -mapas gerados.

A escolha da abertura θ das calotas usadas para calcular os histogramas produz variações nos detalhes dos σ -mapas em escalas muito menores que θ (da ordem de $1/3$).

Este é um efeito esperado, já que o indicador σ é não-local por construção, sendo gerado a partir de todos os pontos da calota. Assim, não podemos tirar conclusões significativas sobre detalhes muito menores que a abertura das calotas (convém lembrar porém que os centros de calotas vizinhas têm separações bem menores que seu raio, da ordem de $1,5^\circ$). A posição do máximo de anisotropia, e a direção do dipolo, não sofrem grandes variações ($< 6^\circ$) para $\theta = 15^\circ, 30^\circ$ ou 45° . Mas verificamos que para $\theta < 15^\circ$ o máximo se torna muito menos significativo; lembramos ainda que raios menores implicam em menos pixels por calotas, o que aumenta o ruído estatístico. O valor escolhido, $\theta = 30^\circ$, é um compromisso entre resolução do σ -mapa, a amplitude da variação de σ e o tempo computacional requerido.

É possível estimar o ruído estatístico associado ao valor de σ , assumindo pequenas flutuações gaussianas (veja Bernui e Villela ([14]) para uma análise detalhada). O seu desvio-padrão $\frac{\Delta\sigma}{\sigma}$ é dado por

$$\frac{\Delta\sigma}{\sigma} = \frac{\sqrt{N_{\text{bins}}}}{n\sqrt{N_{\text{his}}}} \lesssim 0.05, \quad (2.14)$$

onde n é o número de pixels médio por submapa na calota.

Podemos resumir os resultados acima afirmando que os resultados obtidos a partir de um σ -mapa são robustos se $\frac{\Delta\sigma}{\sigma} \lesssim 0.05$. Para os valores escolhidos acima, temos que $\frac{\Delta\sigma}{\sigma} = 0.03$.

Os efeitos do foreground, mesmo após o processo de filtragem, devem ser mais pronunciados próximo ao equador galáctico. Das duas principais regiões de alta anisotropia no σ -mapa, uma se situa próxima ao centro galáctico (e provavelmente tem origem no foreground), e a outra, mais intensa, fica a sudoeste, a cerca de 20° do equador.

Se esta região de alta anisotropia tivesse origem no foreground, ou no método usado para removê-lo, seria de esperar que sua posição ou intensidade mudassem significativamente com o uso de diferentes métodos de filtragem, ou diferentes conjuntos de dados. Para verificar se é este o caso, estudamos os σ -mapas derivados dos outros 6 mapas de RCF mencionados na Seção 2.5.1, que são mostrados na Figura 2.5 (os σ -mapas PPG e COADDED de um ano são omitidos por serem indistinguíveis respectivamente dos mapas LILC e COADDED de 3 anos). É evidente que, a menos de diferença nos detalhes, a proeminente região com alto valor de σ permanece, e é portanto uma característica robusta dos dados colhidos pelo WMAP. A sua presença consistente nos diferentes mapas limpos sugere que ela não tem origem no foreground, ou pelo menos no foreground galáctico, enquanto a sua permanência nos mapas COADDED, e sua pouca variação entre os mapas limpos, indica que ela não é resultado de algum artefato devido ao método de filtragem.

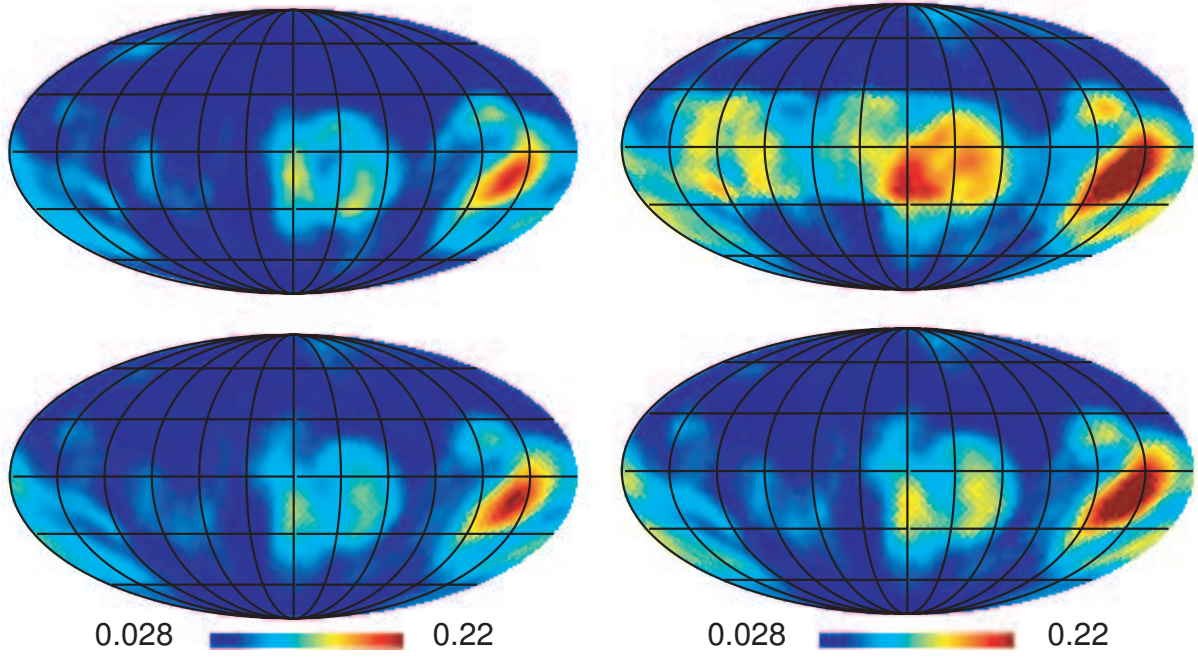


Figura 2.5: Na direção dos ponteiros do relógio estão os seguintes mapas: LILC, COADDED de três anos (mostrando os efeitos da contaminação galáctica não filtrada), OT e TOH. Os mapas COADDED e PPG de um ano são omitidos por serem muito similares aos mapas COADDED de três anos e ILC, respectivamente. Note a presença consistente em todos os mapas da região de alto σ no quadrante sudoeste.

Por outro lado, a outra região de alto sigma observada, situada em torno do centro galáctico, apresenta variações significativas de forma e intensidade entre os vários mapas. De fato, a contaminação galáctica na RCF é mais intensa exatamente na vizinhança do centro da galáxia, onde a emissão e absorção de microondas pela Via Láctea é mais intensa. Assim, seria razoável supor que se os métodos de filtragem não fossem totalmente eficientes em eliminar a contaminação de foreground, então as anisotropias induzidas pela contaminação de origem galáctica se manifestariam com mais intensidade nesta direção. Ao que tudo indica, portanto, o nosso método é sensível a este tipo de contaminação, e consegue determinar a sua direção geral de forma consistente.

2.6 Mapas com máscaras

É bem sabido que em mapas de céu inteiro, mesmo após o processo de filtragem, ainda persiste a presença de contaminação de foreground residual [51], e o ruído tem propriedades estatísticas complexas [78]. O seu uso para estudar estruturas em grandes escalas angulares não é consensual [35], e alguns pesquisadores da área questionam a plausibi-

lidade de um processo de filtragem eliminar quase totalmente a contaminação galáctica sem introduzir novos artefatos observacionais, por mais sofisticado que seja o método usado. Sem entrar no mérito da questão, fizemos também uma análise idêntica [114] à exposta acima, usando o mapa LILC referente a um ano de observações do WMAP, mas aplicando máscaras que eliminam completamente a região em torno do plano galáctico em que a contaminação é mais intensa. Com pequenas modificações, indicadas abaixo, o procedimento para gerar os σ -mapas é idêntico ao descrito acima, então nos limitamos no que se segue a apresentar os resultados, com breves comentários.

Dada a falta de consenso a respeito da utilidade de efetuar uma análise de dados baseados em mapas (limpos) de céu inteiro da RCF, complementamos nossa análise aqui estudando as anisotropias em grandes escalas no mapa LILC após a aplicação da máscara KP2 (o mapa resultante será daqui em diante referido como LILC-KP2), que elimina da análise aproximadamente 15% dos pixels, concentrados principalmente ao longo do plano galáctico.

Para gerar o σ -mapa, precisamos redefinir para cada ponto qual a calota (ou sua equivalente) na qual o valor de σ é calculado. Mas como esta é uma medida não-local, algumas calotas com centro fora da máscara KP2 incluirão regiões que ficam dentro de KP2, e que portanto precisam ser excluídas. Se simplesmente definíssemos em torno de tais pontos uma calota esférica só parcialmente preenchida, obteríamos um valor alto para σ devido à distribuição altamente irregular dos pixels, que seria simplesmente um artefato numérico sem significado observacional. Obviamente, é necessário modificar o método ligeiramente para lidar com mapas com máscaras.

O nosso ponto de partida é o mesmo conjunto de calotas usadas anteriormente, com raio $\theta_0 = 30^\circ$ e regularmente espaçadas. Excluimos as calotas nas quais menos de 15% dos pixels se encontram fora da máscara.

Dentre as calotas restantes, para aquelas que se encontram totalmente fora da máscara, o σ é calculado normalmente.

Para as calotas que intersectam a máscara KPW, calculamos o HSAP levando em conta somente as temperaturas dos pixels fora da máscara. Para obter o HSAP esperado correspondente, geramos 1000 mapas de céu inteiro estatisticamente isotrópicos, e excluimos destes a região coberta por KP2 antes de calcular os histogramas de cada um e o histograma médio. Obtemos então a diferença entre o histograma esperado assim obtido para cada calota, e o histograma observado, para calcular o Υ , e daí o σ . Para um exemplo de histogramas calculados para a mesma calota, com e sem a remoção da máscara, veja a Figura 2.6.

Temos assim um σ -mapa LILC-Kp2, cobrindo todo o céu exceto em pontos situados

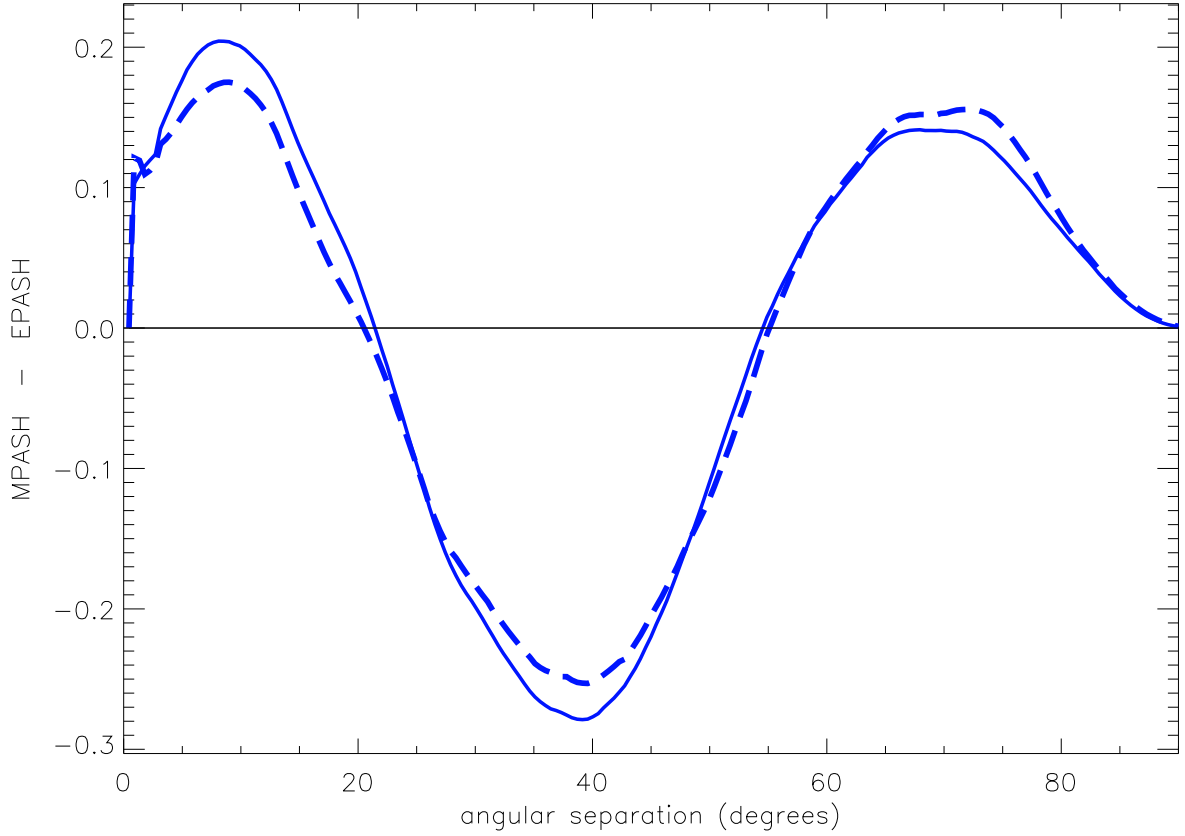


Figura 2.6: O histograma da diferença, Υ , para uma calota centrada no equador galáctico, em $(l, b) = (225^\circ, 90^\circ)$ no mapa LILC, de céu inteiro (linha sólida) e com a máscara KP2 removida (tracejado). Os valores de σ resultantes são praticamente os mesmos: 0.10704 e 0.10061, respectivamente.

na máscara KP2 e adjacências.

Como pode ser observado na Figura 2.7, o eixo de simetria observado nos mapas de céu inteiro também está presente no mapa LILC-KP2, e aponta para aproximadamente a mesma direção. Assim como no caso dos mapas de céu inteiro, o σ -mapa resultante é independente da resolução usada no mapa de temperatura LILC, do número de submapas usados em cada calota, e do número de subdivisões usados nos histogramas. Ele também não depende fortemente (para valores razoáveis) da abertura da calota, θ_0 , embora neste caso estejam presentes distorções próximas aos limites das máscara, que afetam áreas tão maiores quão maior for o valor de θ_0 .

Calculamos em seguida o espectro de potência do mapa LILC-KP2 (veja Figura 2.8). O valor obtido para o dipolo é muito alto, se comparado tanto aos mapas estatisticamente isotrópicos simulados, quanto ao σ -mapa obtido a partir do mesmo mapa sem a aplicação da máscara. Os valores do quadrupolo e octopolo são também altos em relação à media dos mapas simulados, embora o primeiro não seja estatisticamente significativo. Estas características também estão presentes no mapa de céu inteiro LILC, mas os valores exatos

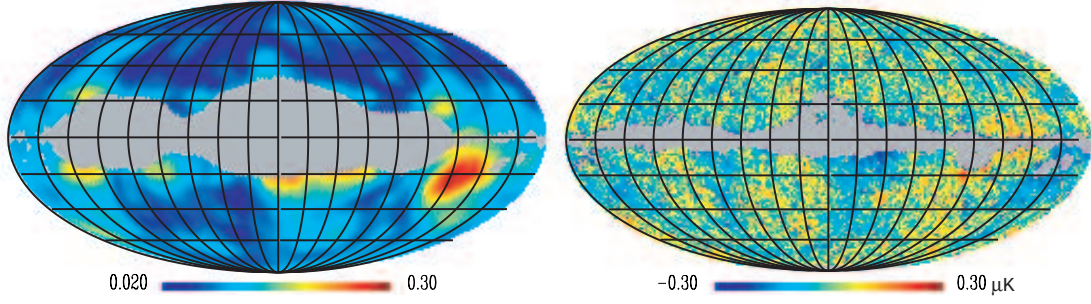


Figura 2.7: O mapa LILC após a aplicação da máscara KP2, e o σ -mapa resultante.

são bastante diferentes. É preciso, porém, ter cautela ao efetuar tais comparações, pois estritamente falando os harmônicos esféricos somente são funções ortogonais se definidos em toda a esfera. O 'espectro de potência' obtido de um mapa após a remoção de uma máscara serve somente para comparações qualitativas, não para uma análise quantitativa precisa.

Assim como fizemos para os mapas de céu inteiro, também calculamos o σ -mapa após a aplicação da máscara KP2 no mapa de RCF Q+V+W Coadded [77]. Cautelosamente, Eriksen *et al.* consideram este o mais confiável dos mapas para estudo da RCF [51, 22]. Da mesma forma que no caso do mapa LILC-KP2, encontramos fundamentalmente a mesma estrutura na anisotropia em grande escala, mais uma vez confirmando a robustez dos nossos resultados.

2.7 O significado das anisotropias

O indicador de anisotropia apresentado acima nos permite mapear direcionalmente os desvios da isotropia estatística no céu de microondas. Aplicando este método a mapas das flutuações de temperatura da RCF, obtemos mapas da anisotropia estatística, os chamados σ -mapas, onde detectamos uma pequena região da esfera celeste onde a anisotropia estatística é muito pronunciada (com alta significação estatística). Esta região essencialmente define um eixo de máxima anisotropia próximo a direções encontradas por métodos completamente diferentes, por Eriksen *et al.* [53] e Land & Magueijo [89].

Resumidamente, para obter uma medida mais quantitativa desta anisotropia, estudamos os coeficientes da expansão em esféricos harmônicos do σ -map gerado a partir do mapa ILC (que procura eliminar as contaminações de foreground). Verificamos que o componente de dipolo correspondente, assim como o quadrupolo e octopolo, são muito maiores do que os valores esperados em mapas estatisticamente isotrópicos. A combinação destes três componentes define uma direção preferencial próxima da região de máximo σ .

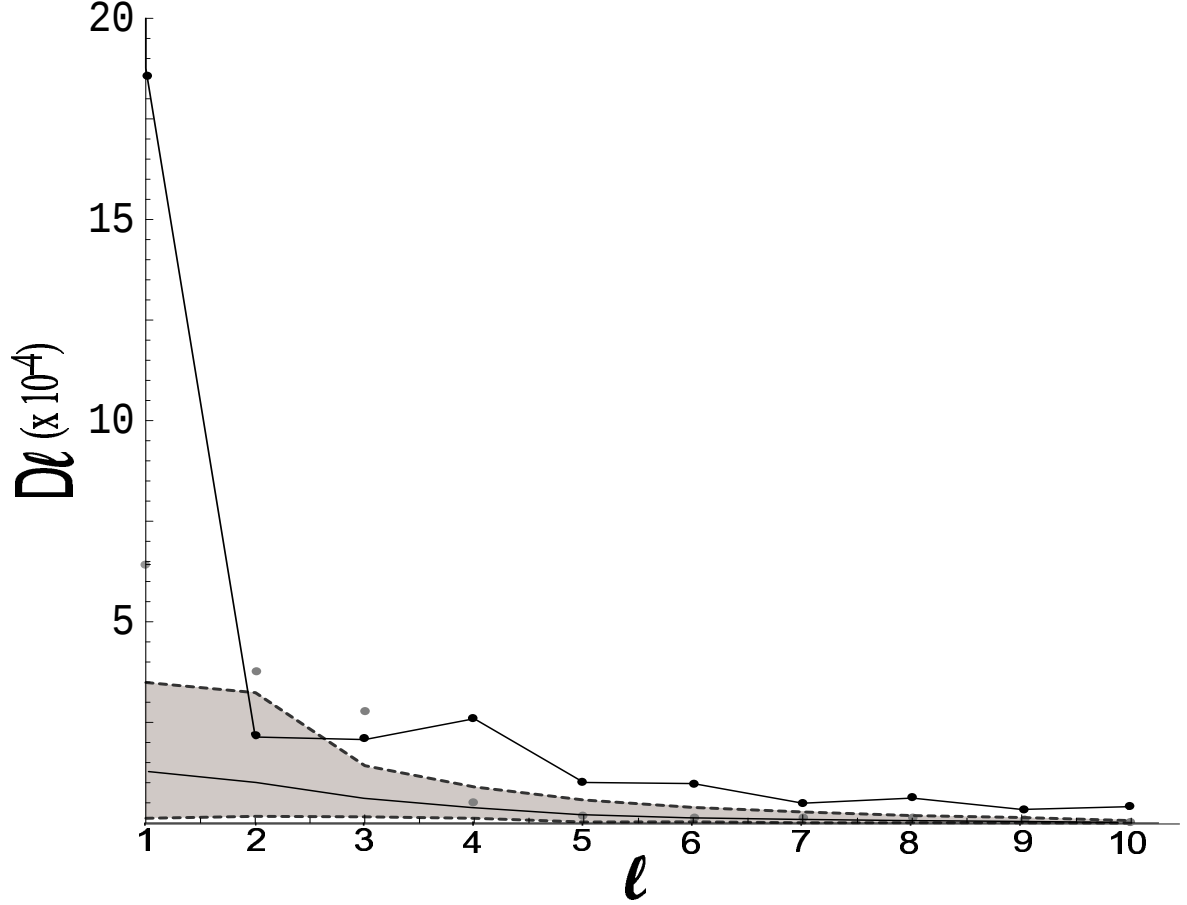


Figura 2.8: O espectro de potência do mapa LILC-KP2 (linha sólida). Para efeitos de comparação, também são mostrados o espectro de potência para o mapa LILC de um ano (círculos em cinza), assim como os valores médios esperados para um mapa de céu inteiro estatisticamente isotrópico, com a concomitante variância cósmica.

Usando um teste de χ^2 mostramos que os altos valores do dipolo, quadrupolo, octopolo e hexadecupolo são estatisticamente significativos, se comparados aos valores esperados no caso estatisticamente isotrópico. Além disso, a probabilidade destes primeiros quatro multipolos terem ao mesmo tempo valores tão altos se o mapa ILC da RCF fosse estatisticamente isotrópico é menor que uma parte em 10^{-6} .

Finalmente, mostramos que os resultados obtidos são robustos no que diz respeito a escolha precisa dos parâmetros para o cálculo de σ , desde que o ruído estatístico seja mantido sob controle. Os resultados são substancialmente os mesmos para um amplo conjunto de mapas de temperatura com diferentes algoritmos de filtragem. Em particular, a região de alto σ no sudoeste do σ -mapa permanece essencialmente invariante para todos os mapas considerados, e o alto valor dos multipolos D_ℓ mais baixos (i.e., baixo ℓ) permanece estatisticamente significativo. Tudo isto indica que a anisotropia detectada é intrínseca aos dados da RCF, e não é um artefato do método de limpeza de foreground

empregado.

O nosso indicador é portanto um instrumento útil, capaz de detectar anisotropias em grandes escalas na RCF que não são imediatamente aparentes a partir da análise direta dos mapas de temperatura. O objetivo, obviamente, não é tornar desnecessário o estudo dos mapas de temperatura, mas sim promover uma abordagem complementar, que revele aspectos diferentes do mesmo conjunto de dados. A relação entre mapas de temperatura e σ -mapas é claramente não trivial; embora uma relação implícita exista, não há uma maneira direta de expressá-la analiticamente, e tal abordagem, mesmo se factível, talvez não produzisse uma fórmula de valor prático.

Em relação à natureza da anisotropia em grande escala revelada no σ -mapa, a questão fundamental que se impõe é: Qual a sua origem? Não há, por enquanto, uma resposta consensual para esta pergunta, mas várias hipóteses têm sido levantadas para explicar anomalias semelhantes (e provavelmente relacionadas) encontradas anteriormente na RCF.

Poderíamos supor que as anomalias tem origem em uma forma particularmente sutil de contaminação de foreground, que não é removida (nem grandemente afetada) por nenhum dos métodos de filtragem analisados (a direção de σ máximo indicaria então a contaminação mais intensa). Neste caso, o nosso entendimento da cosmologia não sofreria grandes abalos, e as anomalias seriam simplesmente consequência da nossa ignorância a respeito da exata composição da vizinhança cósmica imediata. Tal contaminação provavelmente interferiria na detecção de uma possível topologia não trivial do Universo, destruindo, por exemplo, a correlação entre pares de círculos no céu (veja o capítulo 6 para mais detalhes).

A alternativa é supor que tais anomalias tenham origem no fato de que o Universo é genuinamente anisotrópico em grandes escalas. Esta é uma possibilidade instigante, porque questiona uma das hipóteses fundamentais nas quais se baseia a cosmologia atualmente (e portanto potencialmente capaz de alterar o nosso entendimento do Universo). Em particular, os modelos inflacionários usuais precisariam ser modificados, já que prevêem um padrão de flutuações de temperatura da RCF estatisticamente isotrópico.

Talvez a maneira mais natural de produzir direções privilegiadas, e assim quebrar a anisotropia global, é supor que o Universo possua uma topologia não trivial, como explicamos no Capítulo 1 (veja também [35] e [157]). Um estudo detalhado que simule os σ -mapas relacionados a topologias específicas fica além do escopo do presente trabalho (veja o trabalho de Hipolito-Ricaldi & Gomero [79] para um estudo de simulações de tais mapas nos casos estudados no Capítulo 5 para mais detalhes, e também [115]). Se a topologia realmente é a origem da anomalia detectada, então os indicadores de anisotropia que apresentamos podem ser usados para distinguir entre as várias topologias possíveis.

Qualquer que seja a origem destas anomalias em grande escala presentes na RCF, o indicador de anisotropia estatística que descrevemos parece ser robusto e sensível o suficiente para mapeá-las, o que talvez possa nos ajudar a explicá-las.

Capítulo 3

Detectabilidade da topologia cósmica

E aqueles que dançavam eram tidos como loucos por quem não podia ouvir a música.

Friedrich Nietzsche

3.1 Introdução

O recente nascimento da era da cosmologia de precisão [153] tornou possível impor limites extremamente restritivos às características do Universo em que vivemos. Ao que tudo indica, o cosmos é homogêneo em grandes escalas, tem curvatura espacial muito próxima de zero, e a radiação eletromagnética e a matéria bariônica da qual somos formados perfazem somente cerca de 5% do seu conteúdo material. A maior parte da matéria-energia do Universo é aparentemente composta de dois componentes (ou classes de componentes, ou diferentes fases de um mesmo componente), cuja natureza exata desconhecemos. Cerca de um terço da densidade total se deve a um componente sem pressão, capaz de se aglomerar gravitacionalmente de modo a formar a estrutura visível da matéria luminosa em galáxias e aglomerados. Os dois terços restantes correspondem a um componente com pressão negativa que dá origem à fase corrente de expansão acelerada do Universo. Nenhum destes dois componentes é diretamente observável, e são respectivamente denominados, em uma medida semântica da nossa ignorância física, matéria escura e energia escura.

A presença da matéria escura pode ser deduzida por sua influência gravitacional na matéria bariônica visível (curvas de rotação de galáxias, dispersão de velocidades e emissão de raios-X por gás quente em aglomerados), ou por efeitos de lente gravitacional. A

abundância relativa observada dos elementos leves gerados na nucleossíntese primordial mostra que bárions representam no máximo 15% da densidade dos componentes sujeitos a se aglomerar. (c.f. por exemplo [21] e [187]). Análises detalhadas do espectro de potência da matéria da estrutura em larga escala do Universo [165] determinam que a densidade dos componentes sem pressão é cerca de um terço da densidade crítica.

Por outro lado, como a densidade da energia escura tende a se manter uniforme no espaço (a sua natureza repulsiva impede que seja sujeita à instabilidade gravitacional), sua presença só pode ser inferida a partir da observação de efeitos em grandes escalas. A expansão acelerada do Universo, detectada através de observações de supernovas do tipo Ia (SN Ia) como velas-padrão, é a indicação mais clara da sua presença (veja e.g., [133], [119] e [3]). As observações das SNs Ia, em conjunção com o pequeno valor da curvatura indicado pelas medições da radiação cósmica de fundo (RCF, veja a Seção 1.3 e também [158]), indicam que a densidade deste componente é cerca de dois terços da densidade crítica.

Os limites impostos à curvatura seccional resultam em um valor próximo de zero, mas que pode ainda ser negativo, positivo ou nulo. O espaço de cobertura, que define as classes de topologia possíveis, é, portanto, indeterminado. De fato, nenhuma das medições indicadas acima impõe quaisquer restrições à topologia das seções espaciais do Universo (mas restringem a sua detectabilidade, como veremos). Do ponto de vista teórico, não há nenhum modelo comumente aceito que fixe a topologia (mas veja os trabalhos de Brett McInnes [108] para alguns resultados preliminares instigantes). Assim, a topologia cósmica só pode ser atualmente determinada através de observação. Os métodos propostos para detectar e identificar a topologia podem ser divididos em duas classes: Os que se baseiam na detecção indireta de isometrias, e os de detecção direta, a partir da observação da repetição de padrões.

Os métodos indiretos estão fora do escopo desta tese. São potencialmente capazes de detectar, mas não determinar a topologia cósmica. Muito brevemente, eles fazem uso das condições de contorno impostas pelo domínio fundamental para gerar o espectro das flutuações de densidade durante o desacoplamento entre matéria e radiação, para daí obter restrições sobre os multipolos da radiação cósmica de fundo (ou, mais especificamente, sobre o seu espectro angular de potência). Bons exemplos de tal abordagem incluem [79], que faz uso de alguns resultados que apresentamos no Capítulo 5, assim como [142] e [6].

Entre os métodos diretos, um dos mais proeminentes se baseia nos chamados histogramas de separação de pares [92] (HSP, ou PSH no inglês original), que são histogramas da frequência das separações entre pares de objetos discretos em um catálogo; como pares ligados por isometrias translacionais são separados sempre pela mesma distância, se tais

isometrias estivessem presentes, e fossem detectáveis em um dado catálogo, o HSP resultante apresentaria picos bem definidos. Funcionalmente, os HSP são gerados de forma quase idêntica aos HSAP descritos em detalhes no Capítulo 2. Outro método aplicável a fontes discretas, o método de pares correlacionados coletados [173], ou PCC (CCP no original), é capaz de detectar isometrias não translacionais, mas não tem como discriminar entre diferentes topologias. No que se refere a estudos da radiação cósmica de fundo, que correspondem à maior distância que somos capazes de observar diretamente no Universo, temos ainda o método dos chamados círculos no céu [39], que discutimos em maior detalhe no capítulo 6. Rapidamente, a aplicação do grupo de holonomias na superfície de último espalhamento (uma esfera para o observador comóvel) gera imagens, cujas interseções com a própria SUE são círculos. No caso de uma topologia detectável, existirão portanto pares de círculos correlacionados. A sua presença e posição relativa poderiam então ser usadas para determinar a topologia do Universo.

Todos estes métodos exigem, para gerar um resultado positivo, não só a presença de uma topologia cósmica não trivial, mas também a existência de pares de pontos ligados por holonomias que possam ser detectados no tipo de busca efetuado pelo método em questão. Na verdade, estes métodos serão na melhor das hipóteses capazes de detectar somente um subgrupo das holonomias, que não necessariamente será capaz de discriminar entre as várias topologias potenciais. Por tudo isto, determinar a detectabilidade das topologias possíveis é crucial para a interpretação dos resultados obtidos por qualquer um destes métodos. Nas demais seções do presente capítulo, definiremos um indicador de detectabilidade, e usá-lo para estudar sistematicamente a dependência da detectabilidade em termos dos parâmetros observacionais de duas classes de modelos cosmológicos, capazes de reproduzir o comportamento da matéria e energia escuras. Iremos a seguir determinar o que é possível dizer sobre a detectabilidade de classes de topologia de forma independente do método de detecção usado, levando em conta somente os dados observacionais disponíveis no momento. Estes resultados foram apresentados pela primeira vez em [117], [118], [116] e [100].

3.2 Definindo a detectabilidade

De posse de uma classificação completa (nos casos plano e esférico), ou de um conjunto (no caso hiperbólico) das possíveis topologias não-triviais para as seções espaciais M do Universo, um possível próximo passo para um estudo sistemático da topologia cósmica é a determinação das condições de detectabilidade para cada um destes casos, em termos de parâmetros dos modelos cosmológicos. É pouco prático determinar condições suficien-

tes para a detectabilidade, pois estas dependeriam de parâmetros observacionais. Mas é possível, e como veremos, bastante útil, definir de forma sistemática condições necessárias para a detectabilidade (para discussões sobre este tópico com enfoques ligeiramente diferentes, veja [63], [65]).

É importante ressaltar que, estritamente falando, não estamos discutindo a determinação da topologia (o que no caso é equivalente a determinar o seu grupo de holonomias Γ_M), mas tão somente de uma ou mais isometrias que podem ou não ser uma base para todo o grupo. No Capítulo 5 desenvolveremos esta intuição fundamental de que só é possível detectar holonomias específicas associadas a uma dada topologia, para obter conclusões bastante gerais a respeito da chamada forma local do Universo, as informações geométricas e topológicas a que efetivamente temos acesso.

Como vimos no Capítulo 1, em uma seção espacial multiplamente conexa M , existem múltiplas geodésicas fechadas do tipo espaço que passam pela posição de um dado objeto A (digamos, situado em q). No espaço de cobertura, cada uma destas geodésicas corresponde a uma imagem de A gerada pela ação de uma das isometrias γ do grupo de holonomias Γ_M .

Como o espaço de cobertura permanece homogêneo, as posições relativas destas várias imagens permanecem as mesmas, a menos de homotetia (i.e., de variações no fator de escala global). Assim, na variedade pseudo-riemanianna $\mathcal{M}$ dotada da métrica FLRW, do qual M é uma folheação, as linhas de tempo das imagens múltiplas de A seguirão o fluxo de Hubble.

Considere agora um observador comóvel¹ O situado atualmente na Terra, que definimos como a origem do sistema de coordenadas. A linha de tempo do objeto em q nestas coordenadas, parametrizada pelo tempo próprio de O , é dada por $x_q = (t, a(t)q)$, e as linhas de tempo de suas imagens são simplesmente $x_{\gamma q} = (t, a(t)\gamma q) | \forall \gamma \in \Gamma_M$. O observador O será então capaz de observar as imagens múltiplas de A exatamente onde o cone de luz passado de O tiver interseção com as linhas de tempo $x_{\gamma q}$, como pode ser visto na Figura 3.1. Note que as várias imagens, situadas a diferentes distâncias, são observadas em tempos cósmicos distintos. O cone de luz passado pode ser obtido integrando a métrica FLRW (1.10) ao longo de uma geodésica nula ($ds = 0$) radial ($d\theta = d\phi = 0$), de O ($\chi_0 = 0$) até uma das imagens de A . Como fica claro na figura, a distância até cada uma das imagens observadas depende da evolução do fator de escala $a(t)$. Uma escolha mais natural é a definida pelo fator de escala $a(t)$, correspondente a unidades do raio de curvatura para $k \neq 0$. Expressa nestas coordenadas comóveis, a geometria de M não

¹Para uma discussão a respeito dos efeitos do movimento próprio do observador em relação ao referencial comóvel local na detecção da topologia cósmica veja o Capítulo 6.

muda com o tempo (veja novamente a Figura 3.1), e as distâncias entre as imagens, e entre estas e o observador, permanecem constantes. Neste caso, a distância até a origem de um objeto com coordenada radial χ é simplesmente a quantidade adimensional χ .

A condição mínima para que a topologia seja detectada por meio de repetição de padrões é que pelo menos uma das isometrias de seu grupo de holonomias, identificando dois pontos no espaço de cobertura, seja detectável. São necessárias então ao menos duas imagens do mesmo ponto no espaço que estejam situadas no Universo observável (i.e., acessível a observações). Seja $\ell_M(q)$ o comprimento da menor geodésica fechada que passe pelo ponto q (a ação livre das holonomias garante que um comprimento mínimo maior que zero exista para topologias não triviais). Temos

$$\ell_M(q) = \min d(q, \gamma q) \mid \gamma \in \Gamma_M . \quad (3.1)$$

Omitiremos o q daqui para frente, quando não houver possibilidade de confusão. Obviamente para seções espaciais globalmente homogêneas ℓ_M não depende de q . É fácil ver que o chamado raio de injetividade, o raio da maior esfera que pode ser 'inscrita' no domínio fundamental de M , é simplesmente

$$r_{inj} = \frac{\ell_M}{2} . \quad (3.2)$$

Devido às suas propriedades de rigidez, em variedades esféricas e hiperbólicas, r_{inj} (expresso em unidades do raio de curvatura) é um invariante topológico. Variedades planas, por outro lado, não possuem um comprimento característico, e seu r_{inj} pode portanto assumir qualquer valor. Por causa disto, não é possível estabelecer condições gerais sobre a detectabilidade das topologias de variedades planas. Vamos portanto neste capítulo nos ater às variedades não planas.

Claramente, precisamos ser capazes de detectar objetos além do raio de injetividade para detectar uma repetição de padrões de origem topológica. Existem porém limitações tanto instrumentais (sensibilidade dos instrumentos, idade das fontes) quanto cosmológicas (a distância finita do horizonte cosmológico) para a distância comóvel máxima a que temos acesso. Denominaremos $\chi_{obs} = \chi(z_{obs})$ a profundidade máxima para um dado levantamento, onde z_{obs} é o desvio para o vermelho cosmológico máximo no levantamento em questão. Omitimos ainda a dependência dos demais parâmetros do modelo cosmológico. O valor de $\chi(z_{obs})$ depende do método utilizado (e.g., corresponde a $z_{obs} \lesssim 10$ para fontes discreta perenes, tais como galáxias, e $z_{obs} = 1089$ [158] para a radiação cósmica de fundo; o horizonte cosmológico $z \rightarrow \infty$ impõe uma cota superior a χ_{obs}). A Figura 3.1 torna claro que a existência de uma profundidade máxima permite a um observador detectar somente um número finito de imagens de um dado objeto, e

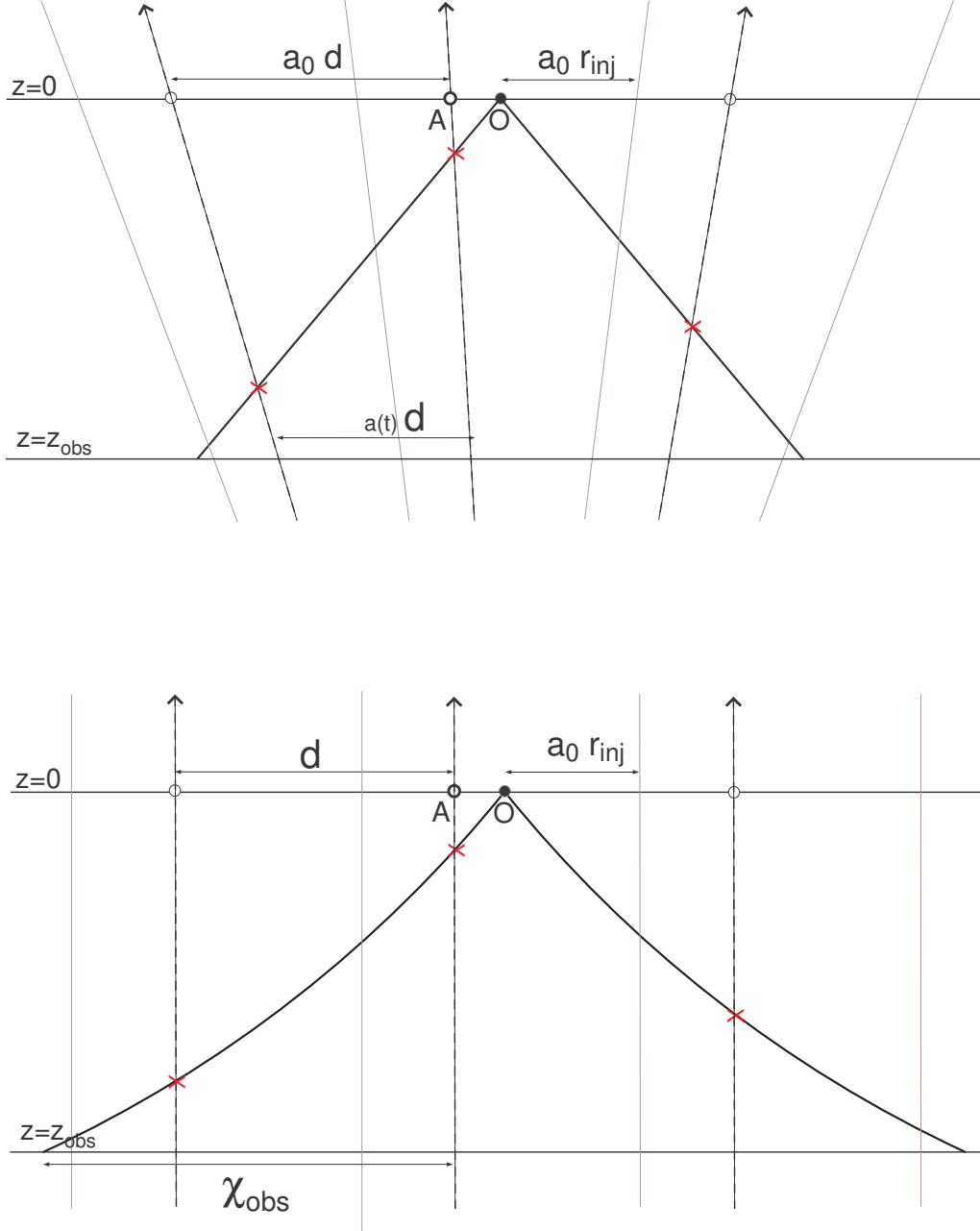


Figura 3.1: Imagens múltiplas (círculos abertos) do objeto A, em repouso em relação ao fluxo de Hubble, que são detectadas por um observador comóvel em O (círculo cheio), em um Universo em expansão. As distâncias são mostradas em unidades métricas (e.g., em Mpc) no painel superior, e em unidades do raio de curvatura no painel inferior. As imagens são observadas em O quando suas linhas de mundo (linhas tracejadas) cruzam o cone de luz passado de O. Os limites do domínio fundamental (linhas cinza) também acompanham o fluxo de Hubble. A profundidade máxima χ_{obs} do levantamento define quais imagens, e portanto isometrias do grupo de holonomias, são detectáveis. Quando expressamos distâncias em unidades do fator de escala, $a(t)$, estas tornam-se independentes do tempo para o observador comóvel. No caso de isometrias em formas espaciais curvas, o fator de escala, que corresponde ao raio de curvatura, define uma escala natural.

como vimos são necessárias pelo menos duas imagens para que a existência de padrões repetidas possa ser aferida.

Podemos assim definir um indicador para detectabilidade para cada variedade não plana como a razão entre o invariante topológico r_{inj} e o parâmetro observacional $\chi(z_{obs})$

$$T_{inj} = \frac{r_{inj}}{\chi_{obs}} . \quad (3.3)$$

Um levantamento no qual $T_{inj} > 1$ não detectará imagens múltiplas de origem topológica, e a topologia cósmica será portanto indetectável por qualquer método de repetição de padrões. Da mesma forma, para $T_{inj} < 1$ a topologia é em princípio detectável, pelo menos para alguns observadores.

Para obter o valor de $\chi(z)$ para um desvio para o vermelho qualquer, é preciso integrar a métrica FLRW (1.10) ao longo de uma geodésica nula ($ds = 0$) radial ($d\theta = d\phi = 0$), de O ($\chi_0 = 0$) até o valor desejado de z (e em particular, até $z_{\max}$ para obter χ_{obs}). Assim, escolhendo novamente unidades tais que $c = 1$, temos

$$\int_t^0 \frac{dt}{a(t)} = \int_0^{\chi(z)} d\chi . \quad (3.4)$$

Lembrando que $(1 + z) = a_0/a(z)$, obtemos

$$\chi(z) = \frac{1}{a_0} \int_0^z \frac{dx}{H(x)} , \quad (3.5)$$

$$\chi(z) = \frac{1}{H_0 a_0} \int_0^z \left[\frac{8\pi G \rho_{ME}}{3H_0^2} + \frac{\Lambda}{3H_0^2} - (1+x)^2 \frac{k}{a_0^2 H_0^2} \right]^{-\frac{1}{2}} dx , \quad (3.6)$$

onde ρ_{ME} é a densidade do conteúdo de matéria-energia do Universo e $H(z) = \dot{a}(t)/a(t)$ é o parâmetro de Hubble. Sabemos por (1.12) que $a_0^2 H_0^2 = k/(1 - \Omega_0)$ (neste caso, $\Omega_0 = \Omega_{ME} + \Omega_{\Lambda 0}$). Usando ainda (1.15a), e a definição dos parâmetros de densidade, escrevemos finalmente a relação desvio para o vermelho—distância em termos do conteúdo material do Universo pela equação

$$\chi(z) = \sqrt{|1 - \Omega_0|} \int_0^z \left[\Omega_{ME} \frac{\rho_0}{\rho_{ME}} + \Omega_{\Lambda 0} - (1+x)^2 (\Omega_0 - 1) \right]^{-\frac{1}{2}} dx , \quad (3.7)$$

Uma característica notável da expressão acima é a sua forte dependência do valor $1 - \Omega_0$ para valores próximos ao caso plano. O valor da distância tende rapidamente a zero quando $\Omega_0 \rightarrow 1$, o que significa que qualquer isometria não-plana se torna indetectável para um valor suficientemente baixo da curvatura. Uma determinação exata dos valores dos componentes de Ω_{ME} e de Ω_{Λ} permitiria estabelecer o espaço de cobertura de M (que depende somente se $k = 0, 1$ ou -1), e também determinar dentre as topologias possíveis

quais têm isometrias detectáveis. Mas qualquer medida observacional possui incertezas inerentes. Em particular, os melhores valores disponíveis dos parâmetros de densidade (e em particular de Ω_0 , c.f. a Seção 1.2.2) admitem seções espaciais localmente planas, esféricas e hiperbólicas. Assim, não é possível descartar a priori a detectabilidade de qualquer topologia específica. Lembrando que modelos inflacionários genéricos prevêm que a diferença $1 - \Omega_0$ deve ser menor por várias ordens de magnitude do que qualquer precisão que poderá ser plausivelmente obtida por experimentos futuros, é possível que mesmo a determinação do espaço de cobertura permaneça em aberto indefinidamente, a não ser que uma topologia não trivial do Universo seja determinada (veja o Capítulo 4 para detalhes de como a determinação da topologia pode restringir parâmetros cosmológicos; e em particular determinar o sinal da curvatura das seções espaciais).

Em face da inevitabilidade das incertezas experimentais, dedicaremos o restante do capítulo à determinação das regiões no espaço paramétrico dos componentes materiais, na qual as várias topologias não-planas são indetectáveis ou potencialmente detectáveis por métodos de repetição de padrões. A idéia fundamental desta análise é que a cada topologia podemos associar um r_{inj} , que define uma curva de contorno que separa o espaço paramétrico nestas duas regiões. Estudaremos em particular duas classes de modelos cosmológicos: Modelos do tipo Λ CDM, onde os únicos componentes significativos são a constante cosmológica e matéria fria (onde predomina a matéria escura, complementada por uma pequena contribuição bariônica), e modelos dominados pelo chamado gás de Chaplygin generalizado, novamente com a adição de um pequeno componente bariônico.

A questão da (in)detectabilidade da topologia foi investigada para valores específicos dos parâmetros de densidade, em [63], [65] e [182]. Os métodos que apresentamos aqui estendem tais análises sistematicamente para todo o espaço paramétrico por meio do estudo das curvas de contorno introduzidos em [117], aplicado a um modelo do tipo Λ CDM, e teve prosseguimento no estudo de modelos dominados pelo gás de Chaplygin generalizado, em [116] e [100]. No capítulo seguinte estudaremos o chamado problema inverso: o que a detecção da topologia potencialmente pode dizer a respeito dos parâmetros dos modelos para o conteúdo de matéria-energia do Universo.

3.3 Modelos Λ CDM: Condições para detectabilidade da topologia

3.3.1 Um modelo simples

Medidas recentes usando supernovas do tipo Ia como velas-padrão [144] indicam fortemente que o Universo esteja há alguns bilhões de anos sofrendo uma expansão acelerada, o que dentro de um contexto FLRW (c.f. (1.11b)) implica na existência de algum componente de energia escura (i.e., com pressão negativa). No modelo Λ CDM esta pressão é produzida por uma constante cosmológica positiva (ou algum componente de energia escura com equação de estado $\rho = -p$). O único outro componente material significativo consiste em matéria fria sem interações significativas que não as gravitacionais, e com parâmetro de densidade Ω_m , que inclui a matéria bariônica. A equação (3.7) tem, neste caso, a forma

$$\chi(z) = \sqrt{|1 - \Omega_0|} \int_0^z [(1+x)^3 \Omega_{m0} + \Omega_{\Lambda 0} - (1+x)^2 (\Omega_0 - 1)]^{-\frac{1}{2}} dx. \quad (3.8)$$

O modelo Λ CDM é atualmente o mais simples compatível com os dados cosmológicos disponíveis, e é por isso muitas vezes usado como o padrão em relação ao qual modelos com componentes de energia escura dinâmicos são comparados.

Para uma dada forma espacial potencial não-plana M , com raio de injetividade r_{inj}^M , tratamos agora de estudar as curvas de contorno no espaço paramétrico $\Omega_{\Lambda 0} - \Omega_{m0}$ definidas por $r_{inj}^M = \chi(z_{obs}, \Omega_{\Lambda 0}, \Omega_{m0})$, que corresponde à seção $\chi(z) = \chi(z_{obs})$ da superfície no espaço $\chi(z) \times \Omega_{m0} \times \Omega_{\Lambda 0}$ (veja a Figura 3.2).

A $\Omega_{\Lambda 0} + \Omega_{m0} = 1$ correspondente ao caso plano (reta plana) pode ser entendida como uma linha de contorno degenerada, que separa o espaço paramétrico em dois semi-planos definidos pelo sinal da curvatura. Cada curva de contorno associada a uma variedade não plana M , definida através da equação $r_{inj}^M = \chi(z_{obs}, \Omega_{\Lambda 0}, \Omega_{m0})$, separa por sua vez um destes semi-planos em duas regiões: Uma situada além da reta plana, na qual a topologia é detectável em princípio, e outra situada entre a curva de contorno e a reta plana, onde a topologia é indetectável.

Outros habitantes notáveis deste espaço paramétrico incluem os pontos de melhor ajuste para os valores dos parâmetros de densidade segundo os dados disponíveis, e os intervalos de confiança ao seu redor (para uma discussão de como tais quantidades são obtidas, veja o Apêndice A). Estas regiões de incerteza definem, para o grau desejado de confiança estatística, onde o nosso Universo 'vive' no plano paramétrico. Uma variedade cuja curva de contorno fique além desta região (em relação à reta plana) é certamente

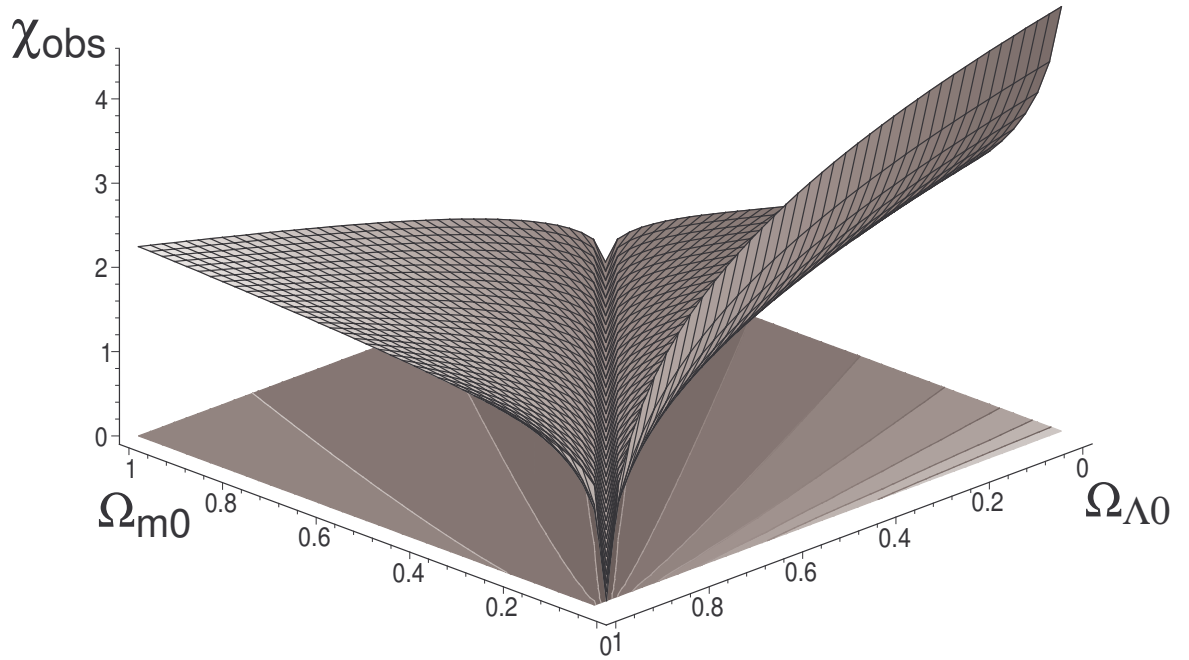
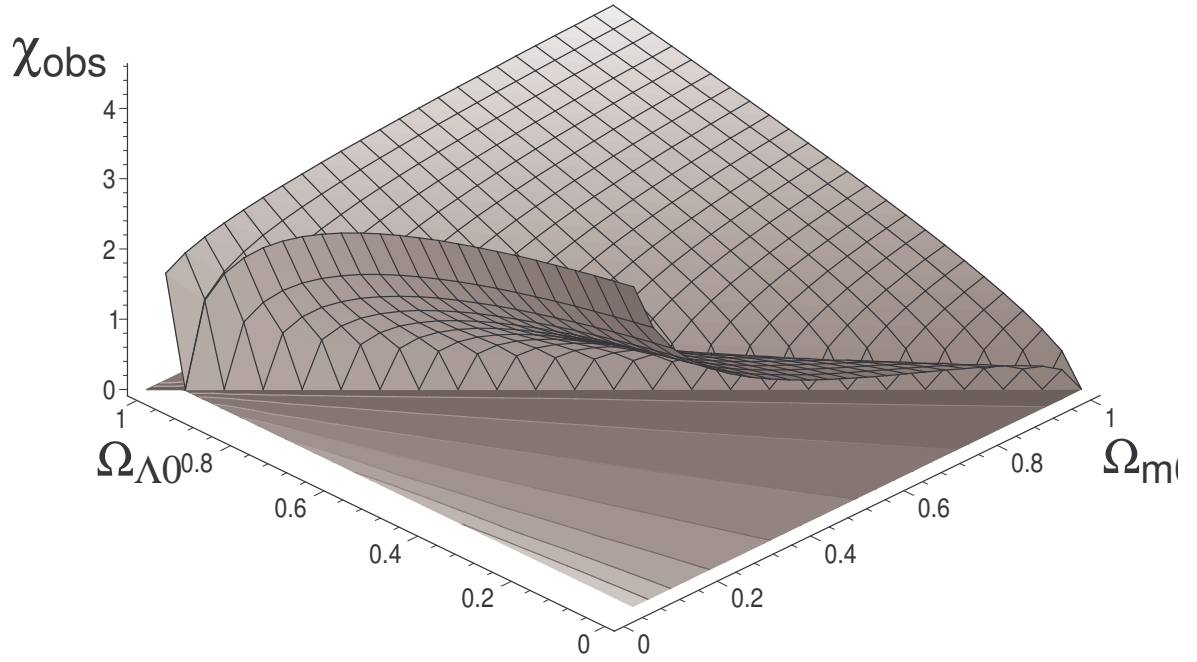


Figura 3.2: Profundidade máxima χ_{obs} de um levantamento em um modelo Λ CDM com $z_{obs} = 1089$, em termos dos parâmetros Ω_{m0} e $\Omega_{\Lambda 0}$. Note a forte sensibilidade ao seu valor exato próximo à reta plana $\Omega_{m0} + \Omega_{\Lambda 0} = 1$. As linhas de contorno, correspondentes às seções $r_{inj}^M = \chi_{obs}$, são projetadas na parte inferior da figura; note o seu formato quase retilíneo, em particular próximo à reta plana.

indetectável por métodos de repetição de padrões. Uma variedade que por outro lado tivesse uma curva de contorno situada entre a região de incerteza e a reta plana seria detectável (em princípio). Mas na realidade a região de incerteza cruza a reta plana, o que significa que nenhuma variedade pode ser considerada definitivamente detectável. Os melhores valores disponíveis atualmente [158] para um modelo Λ CDM, usando as medições feitas pelo satélite WMAP da posição dos picos acústicos da radiação cósmica de fundo, combinadas com medições do parâmetro de Hubble usando supernovas do tipo SN Ia como velas-padrão, produzem os seguintes intervalos de confiança: $\Omega_0 = 1,014 \pm 0,017$ e $\Omega_\Lambda = 0,716 \pm 0,055$, para uma confiança de 2σ (os valores usados na análise em [117] são ligeiramente mais imprecisos, mas não alteram as conclusões finais de forma significativa). Assim, para cada topologia possível, é preciso separar a região de incerteza em sub-regiões onde ela é detectável em princípio, e onde ela é indetectável. Reproduzimos a Figura de [158] (aqui identificada como Figura 3.3), que mostra a região de confiança no plano paramétrico $\Omega_{\Lambda 0} - \Omega_{m0}$ obtida pelo WMAP, sobreposta à sua combinação com uma série de outras observações. Note que em todos os casos curvaturas positivas, negativas e nulas são admissíveis com uma confiança de 1σ . Note ainda que a análise da espectro de potência da radiação cósmica de fundo por si só não é capaz de restringir significativamente todos os parâmetros, apresentando uma considerável degenerescência em uma direção aproximadamente paralela a reta plana. É só combinando estes dados com outras observações que uma região de confiança razoavelmente compacta é obtida.

Uma propriedade importante das curvas de contorno em modelos Λ CDM é a sua convexidade (para $k = 1$) ou concavidade (para $k = -1$) para $z = z_{obs}$ fixo e valores positivos de $\Omega_{\Lambda 0}$ e Ω_{m0} . Isto é bastante evidente na Figura 3.2, e pode ser provado diretamente se considerarmos $\chi_{obs}(z_{obs}, \Omega_{m0}, \Omega_{\Lambda 0}) = r_{inj}^M$ uma função implícita de Ω_0 em função de Ω_{m0} , com z_{obs} e r_{inj}^M como parâmetros. É possível mostrar então que $k \, d^2\Omega_0/d\Omega_{m0}^2 \leq (\geq) 0$ para $k = 1$ ($k = -1$).

Uma análise numérica das curvas de contorno pode ser feita (c.f. o capítulo 4), mas para um tratamento sistemático também é útil usar aproximações mais simples da curva de contorno. Fazemos isto através de dois métodos complementares para obter aproximações lineares das curvas de contorno. O primeiro método, dito da reta tangente, produz condições necessárias, mas não suficientes, de não-detectabilidade de uma dada topologia. O segundo produz condições suficientes mas não necessárias para a não-detectabilidade. Este último método define uma reta secante que pode ser expressa por meio de uma fórmula fechada simples dos parâmetros de densidade. Como veremos, para valores fisicamente relevantes a diferença entre as restrições advindas dos dois métodos é bastante pequena, e portanto é razoável substituir a fórmula analítica da reta secante

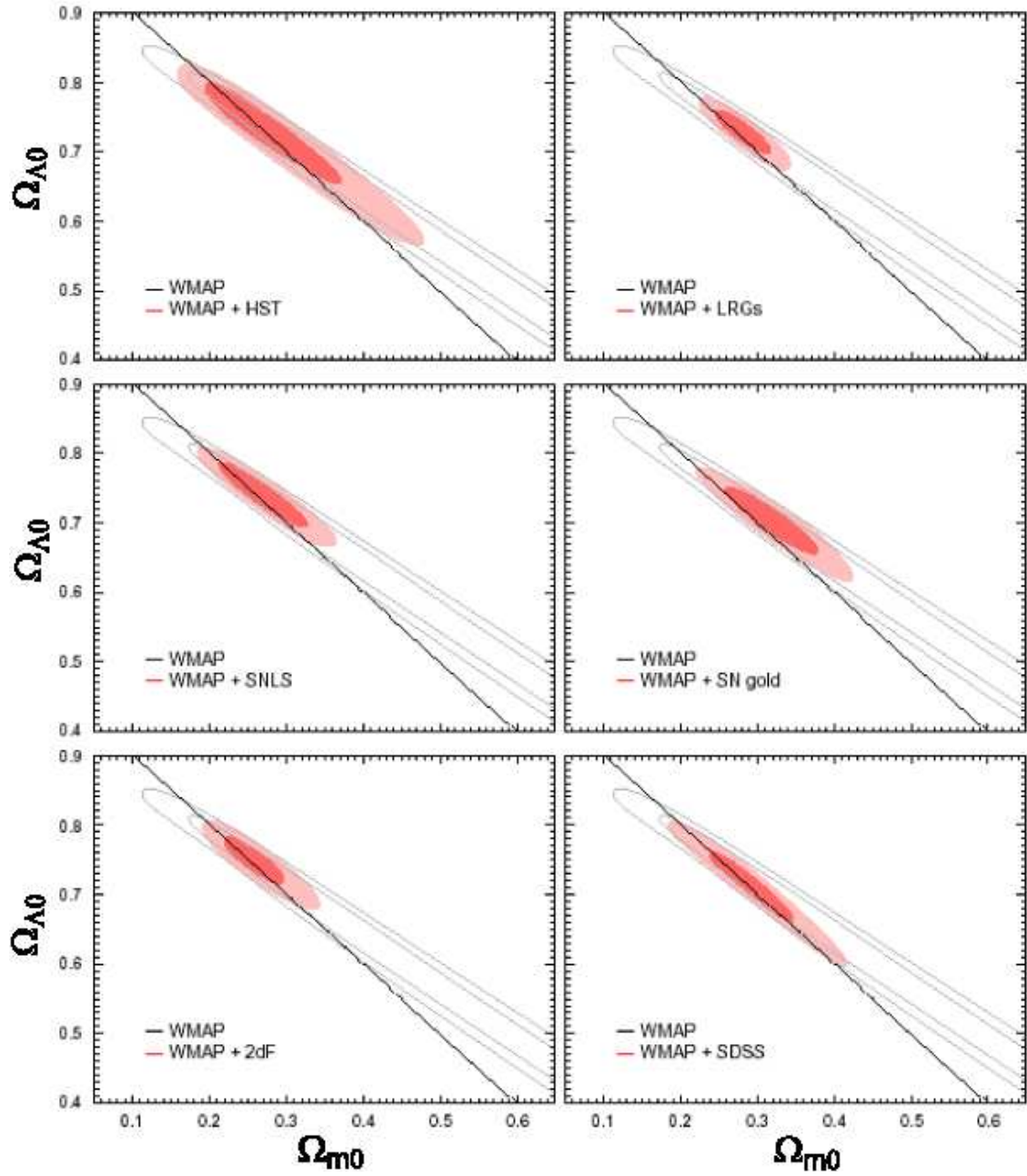


Figura 3.3: Regiões de confiança em 1σ e 2σ obtidas a partir dos dados do WMAP (contornos). A combinação com outros conjuntos de dados observacionais restringe significativamente estas regiões (contornos cheios). Da esquerda para a direita e de cima para baixo, temos, combinados com o WMAP: A medição de H_0 feita pelo telescópio espacial Hubble (HST), e de vínculos advindos de catálogos de galáxias vermelhas luminosas (SDSS LRG), supernovas tipo SN Ia (SNLS, SNGold), galáxias luminosas (2dFGRS) e quasares + galáxias luminosas (SDSS). Fonte: *Spergel et al. (2006)*

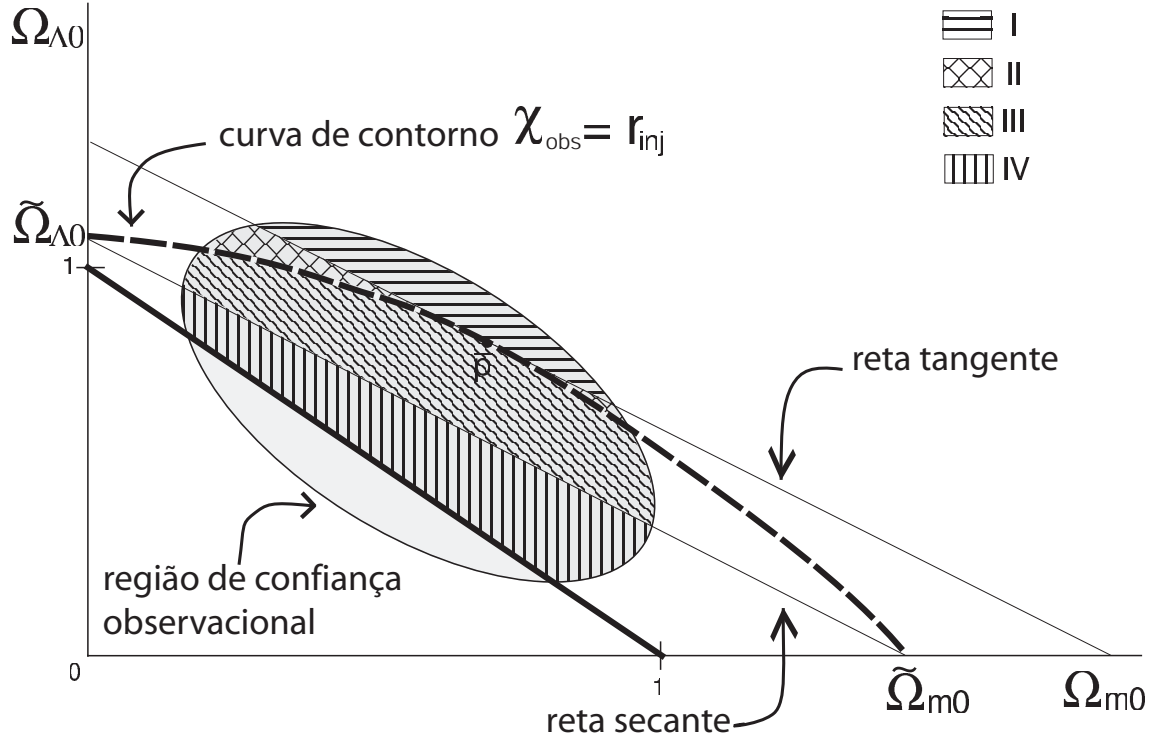


Figura 3.4: Representação esquemática de uma curva de contorno do tipo $\chi_{obs}(z, 0, \tilde{\Omega}_{\Lambda 0}) = r_{inj}$, cruzando a região de confiança para os valores dos parâmetros de densidade. As aproximações das retas secante e tangente são ilustradas. Estas aproximações permitem afirmar que topologia é potencialmente detectável na região I, e indetectável na região IV. As regiões II and III são, respectivamente, detectáveis e indetectáveis, mas não são discriminadas pelas aproximações lineares. O tamanho relativo destas últimas duas regiões em casos fisicamente relevantes é bem menos significativo do que é mostrado na figura.

no lugar da integração forçosamente numérica da equação (3.8). A Figura 3.4 ilustra esquematicamente ambos os métodos e sua relação com a curva de contorno e as regiões de confiança.

3.3.2 Método da reta secante

Considere a curva de contorno $\chi(z_{obs}, \Omega_{m0}, \Omega_{\Lambda 0}) = r_{inj}^M$. Para um dado valor da profundidade do levantamento z_{obs} , definimos a reta secante como sendo a linha que une os pontos $(\tilde{\Omega}_{m0}, 0)$ e $(0, \tilde{\Omega}_{\Lambda 0})$ onde a curva de contorno intercepta respectivamente os eixos Ω_{m0} e $\Omega_{\Lambda 0}$. A equação desta reta é claramente

$$\frac{\Omega_{m0}}{\tilde{\Omega}_{m0}} + \frac{\Omega_{\Lambda 0}}{\tilde{\Omega}_{\Lambda 0}} = 1. \quad (3.9)$$

A grande vantagem advinda do uso da reta secante como aproximação da curva de contorno é que nos pontos de intersecção desta com os eixos a relação de desvio para o

vermelho–distância (3.7) tem um de seus parâmetros zerados, e pode ser integrada analiticamente. Escrevendo explicitamente $\chi(z, \tilde{\Omega}_{m0}, 0) = r_{inj}^M$ e $\chi(z, 0, \tilde{\Omega}_{\Lambda 0}) = r_{inj}^M$, e definindo $\varepsilon = \sqrt{k} = [\sin(1 - \tilde{\Omega}_0)]^{\frac{1}{2}}$ podemos obter respectivamente²,

$$\begin{aligned} r_{inj}^M &= 2\varepsilon \left[\operatorname{arctanh}(1 - \tilde{\Omega}_{m0})^{-1/2} - \operatorname{arctanh}\left(\frac{1 + z \tilde{\Omega}_{m0}}{1 - \tilde{\Omega}_{m0}}\right)^{1/2} \right], \\ r_{inj}^M &= \varepsilon \log \left\{ \frac{(1 + z)(1 - \tilde{\Omega}_{\Lambda 0})^{1/2} + [(1 + z)^2 - z(2 + z)\tilde{\Omega}_{\Lambda 0}]^{1/2}}{1 + (1 - \tilde{\Omega}_{\Lambda 0})^{1/2}} \right\}. \end{aligned} \quad (3.10)$$

Para escrever explicitamente os termos da Equação (3.9) é preciso resolver as equações (3.10) para $\tilde{\Omega}_{m0}$ e $\tilde{\Omega}_{\Lambda 0}$. Embora isto possa ser feito analiticamente, as extensas e complicadas expressões resultantes não são particularmente úteis. Porém, a não ser que exista um horizonte para um valor finito $z_{\max}$ (o que não é o caso para os valores atualmente aceitos para os parâmetros de densidade), $\chi(z)$ cresce monotonicamente com z , e portanto $\chi(\infty)$ é uma cota superior para $\chi(z_{obs})$. Tomando o limite $z_{obs} \rightarrow \infty$ as equações se simplificam bastante. Usando um pouco de álgebra e algumas relações trigonométricas obtemos

$$\begin{aligned} \tilde{\Omega}_{m0} &= \sec^2 \left(\frac{r_{inj}^M}{2\varepsilon} \right), \\ \tilde{\Omega}_{\Lambda 0} &= 1. \end{aligned} \quad (3.11)$$

Devido às propriedades de convexidade (concavidade) discutidas acima, a reta secante não intersepta a curva de contorno para valores positivos de Ω_{m0} e $\Omega_{\Lambda 0}$. Assim, para um dado valor de z_{obs} a desigualdade $\frac{\Omega_{m0}}{\tilde{\Omega}_{m0}} + \frac{\Omega_{\Lambda 0}}{\tilde{\Omega}_{\Lambda 0}} > 1$ (< 1) para $k > 1$ ($k < 1$) é condição suficiente (mas não necessária) para que a topologia não seja detectável. Mas uma topologia não-detectável para $z_{obs} = \infty$ certamente não é detectável para qualquer z_{obs} finito. Portanto, o uso dos coeficientes (3.11) garante que a desigualdade acima é uma condição suficiente para detectabilidade para qualquer valor de z_{obs} . Podemos afirmar portanto que

Um Universo com seção espacial M tem topologia indetectável se

$$\begin{aligned} \cosh^2 \left(\frac{r_{inj}^M}{2} \right) \Omega_{m0} + \Omega_{\Lambda 0} &> 1, \quad \text{para } \Omega_0 < 1, \\ \cos^2 \left(\frac{r_{inj}^M}{2} \right) \Omega_{m0} + \Omega_{\Lambda 0} &< 1, \quad \text{para } \Omega_0 > 1. \end{aligned} \quad (3.12)$$

²As funções $\operatorname{arctanh}$ e $\log$ aqui são definidas como as extensões analíticas para o plano complexo das funções reais ordinárias, e portanto podem ser aplicadas tanto para casos esféricos como hiperbólicos.

onde r_{inj}^M é o raio de injetividade da variedade M .

Duas outras formas equivalentes, e ocasionalmente mais convenientes, de expressar as desigualdades acima em termos de Ω_0 , são:

$$\begin{aligned}\Omega_0 &> \operatorname{sech}^2(r_{inj}^M/2) + \tanh^2(r_{inj}^M/2) \Omega_{\Lambda 0}, \quad \text{para } \Omega_0 < 1, \\ \Omega_0 &< \sec^2(r_{inj}^M/2) - \tan^2(r_{inj}^M/2) \Omega_{\Lambda 0}, \quad \text{para } \Omega_0 > 1,\end{aligned}\quad (3.13)$$

e

$$\begin{aligned}\Omega_0 &> 1 - \sinh^2(r_{inj}^M/2) \Omega_{m0}, \quad \text{para } \Omega_0 < 1, \\ \Omega_0 &< 1 + \sin^2(r_{inj}^M/2) \Omega_{m0}, \quad \text{para } \Omega_0 > 1.\end{aligned}\quad (3.14)$$

As desigualdades acima têm a virtude da simplicidade, e podem testar a detectabilidade da topologia para *qualquer* valor de z em todo o plano paramétrico. A exigência de que $z_{obs} \rightarrow \infty$ não é significativamente mais restritiva do que o valor $z_{obs} = 1089$ usado em levantamentos da radiação cósmica de fundo (c.f. Capítulo 2).

Usaremos a seguir a reta secante descrita por (3.12) como uma aproximação analítica da curva de contorno. Mas antes é preciso mostrar que ela é uma boa aproximação. Vimos que a reta secante fornece uma condição necessária para detectabilidade da topologia. O que faremos a seguir é apresentar uma outra aproximação linear que fornece uma condição suficiente para tanto, e mostrar numericamente que estas duas condições, uma necessária e outra suficiente, não diferem significativamente uma da outra, sendo portanto boas aproximações para a curva de contorno.

3.3.3 Método da reta tangente

A reta tangente a uma dada curva de contorno é por definição a melhor aproximação linear da curva em uma vizinhança do ponto de contato. Em qualquer ponto $\bar{P} = (\bar{\Omega}_{m0}, \bar{\Omega}_{\Lambda 0})$ de uma curva de contorno a dois parâmetros como $\chi_{obs}(\Omega_{m0}, \Omega_{\Lambda 0}) = r_{inj}^M$, a reta tangente será perpendicular ao gradiente $\nabla \chi_{obs}(z_{obs}, \Omega_{m0}, \Omega_{\Lambda 0})$. A equação desta reta é portanto

$$\frac{\Omega_{\Lambda 0} - \bar{\Omega}_{\Lambda 0}}{\Omega_{m0} - \bar{\Omega}_{m0}} = - \frac{\left. \frac{\partial \chi}{\partial \Omega_{m0}} \right|_{P=\bar{P}}}{\left. \frac{\partial \chi}{\partial \Omega_{\Lambda 0}} \right|_{P=\bar{P}}} . \quad (3.15)$$

Rearranjando os termos de (3.15) temos

$$\alpha \Omega_{m0} + \beta \Omega_{\Lambda 0} = 1 , \quad (3.16)$$

$$\alpha = \frac{\frac{\partial \chi}{\partial \Omega_{m0}}}{\frac{\partial \chi}{\partial \Omega_{m0}} \bar{\Omega}_{m0} + \frac{\partial \chi}{\partial \Omega_{\Lambda 0}} \bar{\Omega}_{\Lambda 0}} ,$$

$$\beta = \frac{\frac{\partial \chi}{\partial \Omega_{\Lambda 0}}}{\frac{\partial \chi}{\partial \Omega_{m0}} \bar{\Omega}_{m0} + \frac{\partial \chi}{\partial \Omega_{\Lambda 0}} \bar{\Omega}_{\Lambda 0}} ,$$

onde omitimos para maior clareza a indicação de que todas as derivadas parciais devem ser tomadas em $P = \bar{P}$.

Note que a propriedade de convexidade (concavidade) mencionada acima garante que a curva nunca cruzará a reta tangente. Assim, para um dado z_{obs} , a condição $\alpha \Omega_{m0} + \beta \Omega_{\Lambda 0} < 1$ (> 1) para $k > 1$ ($k < 1$) é, como afirmamos, necessária mas não suficiente, para garantir a não-detectabilidade da topologia. Trocando o sinal da desigualdade, temos uma condição suficiente mas não necessária para detectabilidade em princípio.

A escolha exata de $\bar{P}$ não é particularmente importante, mas podemos tomá-lo como o ponto da curva de contorno situado ao longo da direção do gradiente $\nabla \chi(z_{obs})$ calculado no ponto de maior $\chi(z_{obs})$ situado dentro da região de confiança. Por depender da escolha do ponto tangente, a reta tangente não apresenta grande interesse intrínseco, mas a usaremos para mostrar que para variedades de interesse a reta secante é uma boa aproximação da curva de contorno.

3.3.4 As curvas de contorno e suas aproximações

As relações (3.12) - (3.14) dependem de dois parâmetros. Mas próximo à reta plana, o parâmetro mais significativo é sem dúvida Ω_0 , e é portanto mais prático utilizar (3.13) ou (3.14) ao invés de (3.12). Abaixo, mostraremos que as retas secante e tangente (que têm entre elas a curva de contorno) são bastante próximas uma da outra para variedades não-planas de interesse. Faremos isto inicialmente deixando que ambos os parâmetros variem, e em seguida fixaremos o menos importante com um valor fisicamente plausível. São comparações simples que podem ser feitas de forma similar para variedades esféricas e hiperbólicas; mas para evitar repetições vamos nos limitar aqui a tabular os resultados das comparações a dois parâmetros para espaços hiperbólicos, e os resultados a um parâmetro para variedades esféricas. A escolha das variedades, além do seu interesse intrínseco, também foi feita para tornar possível a comparação com resultados similares apresentados anteriormente na literatura.

Uma aproximação linear de uma curva de contorno de χ_{obs} tem a forma geral $\Omega_0 = K + \gamma \Omega_{m0}$. Tabulamos abaixo os coeficientes K e γ para os 7 espaços hiperbólicos de menor

volume no censo Hodgson-Weeks [80], calculados para a reta tangente e a reta secante para $z_{obs} = 1089$ (correspondente à última superfície de espalhamento, c.f. [158]), e para a reta secante para $z_{obs} = \infty$. Compare com os resultados para as mesmas variedades, mas para o par de parâmetros $(\Omega_0, \Omega_\Lambda)$, obtidos em [63] para alguns valores específicos e em [117] para todo o espaço paramétrico.

Variedade	r_{inj}	Método RT $z = 1100$		Método RS $z = 1100$		Método RS $z \rightarrow \infty$	
		K	γ	K	γ	K	γ
$m004(1, 2)$	0.183	0.999	-0.008	1.000	-0.009	1.000	-0.008
$m004(6, 1)$	0.240	0.999	-0.014	1.000	-0.015	1.000	-0.014
$m003(-4, 3)$	0.288	0.998	-0.020	1.000	-0.022	1.000	-0.020
$m003(-2, 3)$	0.289	0.998	-0.020	1.000	-0.022	1.000	-0.020
$m003(-3, 1)$	0.292	0.997	-0.020	1.000	-0.022	1.000	-0.022
$m009(4, 1)$	0.397	0.996	-0.038	0.999	-0.042	1.000	-0.041
$m007(3, 1)$	0.416	0.995	-0.042	1.001	-0.047	1.000	-0.043

Tabela 3.1: Coeficientes K e γ para as aproximações lineares com a forma $\Omega_0 = K + \gamma\Omega_{m0}$ para curvas de contorno das 7 3-variedades hiperbólicas, compactas e orientáveis conhecidas de menor volume, ordenadas por r_{inj} crescente. Foram usadas as aproximações da reta tangente (RT) e reta secante (RS) para $z_{obs} = 1089$, e reta secante para $z_{obs} = \infty$. Note que os coeficientes são muito próximos para todas as aproximações em todas as variedades.

A Tabela 3.1 mostra claramente que para todas as variedades estudadas os coeficientes para as várias aproximações lineares são muito próximos, pelo menos para buscas de repetição de padrões na radiação cósmica de fundo. Exatamente o mesmo resultado advém do estudo da detectabilidade de variedades esféricas com raios de injetividade similares. Para todos os efeitos práticos, portanto, a relação (3.14) (ou suas equivalentes) é uma aproximação excelente das curvas de contorno do tipo $\chi_{obs} = r_{inj}^M$ que cruzam a região de confiança.

É porém importante notar que como as holonomias das 3-variedades hiperbólicas multiplamente conexas não incluem translações de Clifford, nenhuma destas variedades é globalmente homogênea. Assim, embora a sua eventual não-detectabilidade seja uma propriedade global, a detectabilidade, mesmo em princípio, depende da posição do observador. Weeks estudou numericamente [182] os *perfis de injetividade* de algumas 3-variedades hiperbólicas, i.e., a fração do volume de cada variedade na qual a topologia é detectável para um dado χ_{obs} . Formalmente, definimos o chamado raio de injetividade local, $r_{inj}^M(x)$,

como metade do comprimento da menor geodésica fechada que passa pelo ponto x . Então, se para uma dada distância r , $v(r)$ é a fração da variedade na qual $\chi_{obs} \geq r$, o perfil de injetividade de M é definido como uma distribuição do tipo $\rho(r_{inj}) = \frac{dv(r_{inj})}{dr}$. Um estudo similar, mas desta vez analítico, pode ser feito para variedades planas [113].

Ainda a partir da Tabela 3.1, fica claro que para valores fisicamente relevantes χ_{obs} não depende fortemente do valor de Ω_{m0} ou de $\Omega_{\Lambda0}$, e pode portanto ser visto aproximadamente como uma função de Ω_0 que tem Ω_{m0} (ou $\Omega_{\Lambda0}$) como parâmetro fixo. Além disso, vimos que a reta secante no limite $z \rightarrow \infty$ é uma boa aproximação das curvas de contorno potencialmente detectáveis em variedades quase planas. Expresso desta forma, (3.13) pode ser entendido como uma relação analítica entre χ_{obs} e Ω_0 que pode ser usada para investigar as propriedades de detectabilidade de amplas classes de formas espaciais não-planas. Mostraremos um exemplo de uma aplicação mais adiante. Uma outra aplicação, para espaços esféricos com restrições ao valor da curvatura muito mais estrito do que as usadas aqui, pode ser encontrada no Capítulo 5.

Passamos agora para as variedades esféricas. Na Tabela 3.2 mostramos as cotas superior e inferior para o valor mínimo de Ω_0 que garante a indetectabilidade da topologia das variedades esféricas globalmente homogêneas (que são exatamente aquelas ditas de ação simples, veja o Apêndice B), para um valor fixo de Ω_{m0} . Estas são as variedades não-planas de detecção potencialmente mais simples. De fato, o método mais bem estudado de procura por repetição de padrões para fontes discretas, baseado em HSPs, só é sensível a isometrias translacionais (para uma prova, veja [60]), enquanto o método de círculos no céu, usado em mapas de radiação cósmica de fundo, se torna computacionalmente muito mais simples no mesmo caso. É importante ainda notar que em todas as variedades esféricas não-globalmente homogêneas que admitem translações de Clifford como holonomias, estas são sempre geradas por um subgrupo translacional que corresponde a um dos casos tabelados. Nestes casos, as cotas aqui obtidas ainda são aplicáveis, mas se referem à menor separação entre um par de pontos no espaço ligados por uma holonomia translacional. Posto de outra maneira, o valor de r_{inj} associado à um grupo finito Γ_T de holonomias translacionais de $\mathbb{S}^3$ é uma cota superior para r_{inj} em variedades com um grupo de holonomias Γ tal que $\Gamma_T \subset \Gamma$. No Capítulo 5 apresentamos uma análise da estrutura dos grupos de holonomias de variedades esféricas, e uma caracterização completa dos seus subgrupos translacionais.

Voltamos agora nosso foco às variedades globalmente homogêneas. Algumas delas foram tratadas isoladamente em [182]; para permitir comparações diretas tomamos o mesmo valor para o parâmetro de densidade da matéria, $\Omega_{m0} = 0,35$. Como veremos, porém, que uma variação no valor de Ω_{m0} dentro de uma faixa observacionalmente razoável

não traz grandes alterações para as condições de indetectabilidade.

Fica claro nas Tabelas 3.2 e 3.3 que as condições obtidas pelos métodos das retas secante e tangente são bastante próximas, e também são próximas dos valores apresentados em [182]. Tabelaamos também os resultados para alguns espaços-lente e diédricos binários selecionados como exemplos.

Variedade	r_{inj}	Sup. Ω_0 Método RT $z = 1100$	Inf. Ω_0 Método RS $z \rightarrow \infty$
S^3/I^*	$\frac{\pi}{10}$	1.011	1.009
S^3/O^*	$\frac{\pi}{8}$	1.018	1.013
S^3/T^*	$\frac{\pi}{6}$	1.031	1.023
S^3/D_m^*	$\frac{\pi}{2m}$	—	$1 + \sin^2\left(\frac{\pi}{4m}\right) \Omega_{m0}$
S^3/Z_p	$\frac{\pi}{p}$	—	$1 + \sin^2\left(\frac{\pi}{2p}\right) \Omega_{m0}$

Tabela 3.2: Cotas superior e inferior para o valor mínimo de Ω_0 que garante a indetectabilidade da topologia das variedades esféricas globalmente homogêneas, com $\Omega_{m0} = 0,35$ fixo. São mostrados para fins comparativos as cotas obtidas pelos métodos tangente e secante para $z_{obs} = 1089$, e para o método secante para $z_{obs} = \infty$. Espaços lentes $L(p, q)$ são globalmente homogêneos somente para $q = 1$, mas o valor $r_{inj} = \frac{\pi}{p}$ é o mesmo para qualquer q .

m	p	r_{inj}	Max. Ω_0 Método RT $z = 1100$
2	4	0.785	1.072
3	6	0.524	1.031
4	8	0.393	1.018
5	10	0.314	1.011
6	12	0.262	1.008
9	18	0.175	1.005

Tabela 3.3: Cotas superiores de Ω_0 para a indetectabilidade da topologia cósmica obtidas a partir do método da reta tangente, para alguns espaços diédricos binários D_m^* e cíclico Z_p , para $\Omega_{m0} = 0,35$. Tais grupos cíclicos são holonomias de espaços-lente lente $L(p, 1)$, mas as cotas aqui obtidas se aplicam a todos os espaços-lente lente $L(p, q)$.

Uma outra maneira de obter cotas para o valor de Ω_0 , sem definir a priori um valor para Ω_{m0} , é sugerido pela forma de elipsóide alongado que toma a região de confiança quando somente dados do WMAP são levados em conta, evidente na Figura 3.3. A reta

na qual se situa o semi-eixo maior desta região tem seu melhor ajuste (veja [158], pp. 58) dado pela equação $\Omega_0 = 1,306 - 0,876 \Omega_{m0}$. Obviamente este é um ajuste observacional, que varia para o conjunto de observações usados. Mas tomando esta reta como um vínculo variável para Ω_{m0} (ao invés de um vínculo fixo, $\Omega_{m0} = \text{const.}$), é fácil ver que ela é dividida por cada curva de contorno em duas semi-retas nas quais a topologia é respectivamente detectável em princípio e indetectável. Usando a aproximação da reta secante para as curvas de contorno, podemos obter a seguinte condição de indetectabilidade:

Um Universo com seção espacial M tem topologia indetectável se

$$\begin{aligned} \Omega_0 &> \frac{1 - 0,766 \sinh^2(r_{inj}^M / 2)}{1 + 1,490 \sinh^2(r_{inj}^M / 2)}, \quad \text{para } \Omega_0 < 1, \\ \Omega_0 &< \frac{1 + 0,766 \sin^2(r_{inj}^M / 2)}{1 - 1,490 \sin^2(r_{inj}^M / 2)}, \quad \text{para } \Omega_0 > 1. \end{aligned} \quad (3.17)$$

Obtemos assim uma boa aproximação para as condições de detectabilidade (em princípio) para as variedades localmente esféricas e hiperbólicas estudadas nas Tabelas 3.1 e 3.2, em termos unicamente de Ω_0 . Os resultados são mostrados na Tabela 3.4. Compare os vínculos obtidos com as condições correspondentes obtidas a partir de um modelo do tipo gás de Chaplygin generalizado, na Tabela 3.5.

Variedade	r_{inj}	Ω_0 min	Variedade	r_{inj}	Ω_0 max
Esférica			Hiperbólica		
$S^3/D_9^*, S^3/Z_{18}$	$\frac{\pi}{18}$	1.006	m004(1,2)	0.183	0.994
$S^3/D_6^*, S^3/Z_{12}$	$\frac{\pi}{12}$	1.013	m004(6,1)	0.240	0.989
$S^3/I^*, S^3/D_5^*, S^3/Z_{10}$	$\frac{\pi}{10}$	1.019	m003(-4,3)	0.288	0.984
$S^3/O^*, S^3/D_4^*, S^3/Z_8$	$\frac{\pi}{8}$	1.030	m003(-2,3)	0.289	0.984
$S^3/T^*, S^3/D_3^*, S^3/Z_6$	$\frac{\pi}{6}$	1.057	m003(-3,1)	0.292	0.984
$S^3/D_2^*, S^3/Z_4$	$\frac{\pi}{4}$	1.143	m009(4,1)	0.387	0.972
			m007(3,1)	0.416	0.968

Tabela 3.4: Valores mínimos (máximos) para Ω_0 que garantem detectabilidade em princípio em pelo menos algum ponto de algumas variedades esféricas (hiperbólicas), assumindo o vínculo $\Omega_0 = 1.306 - 0.876 \Omega_{m0}$, que corresponde à direção de máxima degenerescência da região de confiança obtida a partir de dados do WMAP, assumindo um modelo Λ CDM.

Finalmente, é interessante usar a expressão analítica (3.14) para gerar uma condição suficiente para indetectabilidade em famílias dos espaços diédricos binários e lentes, já que

a expressão em questão depende explicitamente de Ω_{m0} . Para tais variedades as condições de indetectabilidade podem ser dadas de forma geral como funções de, respectivamente, p e m (espaços-lente $L(p, q)$ são globalmente homogêneos para $q = 1$, mas o valor $r_{inj} = \frac{\pi}{p}$ é o mesmo para qualquer q). Como vimos, para qualquer $\Omega_0 > 1$, existem sempre infinitos elementos de cada classe que permanecem detectáveis em princípio. Podemos então definir m_* (ou p_*) como o valor máximo de m (ou p) tal que a topologia com (sub)grupo de holonomias D_m^* (ou Z_p , e por extensão $Z_{p,q}$) é indetectável. Resolvendo a equação (3.14) para m_* e p_* obtemos as condições gerais

$$\begin{aligned} p_* &= \text{Int} \left\{ \frac{\pi}{2} \left[\arcsin \sqrt{\frac{\Omega_0 - 1}{\Omega_{m0}}} \right]^{-1} \right\}, \\ m_* &= \text{Int} \left\{ \frac{\pi}{4} \left[\arcsin \sqrt{\frac{\Omega_0 - 1}{\Omega_{m0}}} \right]^{-1} \right\}, \end{aligned} \quad (3.18)$$

onde $\text{Int}[x]$ denota a parte inteira de x . Os resultados acima generalizam aqueles obtidos em [63] para alguns casos específicos.

3.4 O Gás de Chaplygin generalizado

3.4.1 Um fluido único

Embora um Universo dotado de uma constante cosmológica positiva e preenchido por matéria escura fria, e descrito portanto por modelos Λ CDM, esteja em acordo com as observações, outras alternativas existem que também despertam considerável interesse. Em particular, é interessante presumir que matéria escura e energia escura, ao invés de serem dois componentes materiais distintos, sejam simplesmente manifestações diferentes de um fluido único. Tal abordagem tem a virtude de reduzir a maior parte do conteúdo material do Universo a somente um componente; o preço a pagar é uma equação de estado mais complicada associada a este componente. Fundamentalmente, este fluido precisa ser tal que em desvios para o vermelho grandes, ou no interior de aglomerados de galáxias, ele se comporte como matéria escura e se aglomere de maneira a formar estruturas tais como as observadas atualmente, enquanto em baixas densidades (tais como a densidade média do Universo próximo à época atual, $z \lesssim 1$) a sua pressão se torna cada vez mais negativa, para explicar a atual fase de expansão acelerada do Universo.

Candidatos para tal componente único incluem o gás de Chaplygin (GC), e suas várias generalizações (uma abordagem didática e completa pode ser encontrada em [99]; veja

também [83], [12] e [23]). No chamado *gas de Chaplygin generalizado* (GCG, c.f. [99]), pressão e densidade são relacionadas pela equação de estado

$$p_{\text{Ch}} = -\frac{M^{4(\alpha+1)}}{\rho_{\text{Ch}}^\alpha}, \quad (3.19)$$

onde M é uma escala com dimensão de massa e α é um número real adimensional. O GC corresponde ao caso $\alpha = 1$.

A lei de potência (3.19) garante que para $\alpha > 0$ o GCG tem pressão muito pequena ($|p| \ll \rho$, comparável à $p_m = 0$ para 'poeira') quando a densidade é alta, e também gera uma pressão negativa significativa para baixas densidades. Lembramos que pela segunda equação de Friedmann, (1.11b), para $\Lambda = 0$ temos que $\ddot{a} = -4\pi G(\rho + 3p)/3$. Assim, a escala M determina para qual valor de densidade se dá a transição entre um GCG atrativo e um repulsivo ($\rho_{\text{Ch}}^T = 3^{\frac{1}{\alpha+1}} M^4$), enquanto o parâmetro α determina a rapidez com a qual esta transição ocorre ($\left. \frac{dp_{\text{Ch}}}{d\rho_{\text{Ch}}} \right|_{\rho_{\text{Ch}}=\rho_{\text{Ch}}^T} = \frac{\alpha}{3}$).

Considere agora a geometria FLRW de fundo de um Universo dominado por um GCG. Neste caso, como densidade e pressão médias não dependem da posição, podemos resolver a equação de conservação local de energia para um fluido perfeito (1.13) para obter

$$\rho_{\text{Ch}} = \rho_{\text{Ch}0} \left[(1 - A) \left(\frac{a_0}{a} \right)^{3(\alpha+1)} + A \right]^{1/(\alpha+1)}, \quad (3.20)$$

onde $A = (M^4/\rho_{\text{Ch}0})^{(\alpha+1)}$ e o subscrito "0" denota valores tomados na época atual. Para garantir a positividade de ρ (a condição de energia fraca), tomamos $A \geq 0$. Tomaremos como hipótese também $A \leq 1$, que garante que a pressão não será divergente no futuro. Podemos ainda combinar as duas equações acima para obter $p_{\text{Ch}}(a)$, o que nos permite concluir que para que a fase de aceleração seja posterior à fase de desaceleração em um Universo em expansão (i.e., $a_0 > a$), é necessário que p_{Ch} seja uma função monotonicamente decrescente de a , o que se dá se e somente se $\alpha > -1$, que tomaremos como hipótese.

Como consequência desta última hipótese, quando $a/a_0 \ll 1$, então $\rho_{\text{Ch}0} \propto a^{-3}$, e o GCG se comporta como matéria escura. Posteriormente, quando $a/a_0 \gg 1$, temos no limite $p_{\text{Ch}0} = -\rho_{\text{Ch}0} = \text{const.}$, correspondente a uma constante cosmológica positiva com valor $\Lambda = 8\pi G M^4$. Assim, o GCG se assemelha sucessivamente a matéria escura e a energia escura, à medida que o Universo se expande e se torna menos denso. Temos como casos especiais do GCG o gás de Chaplygin clássico, para $\alpha = 1$, e o modelo Λ CDM, para $\alpha = 0$.

Um modelo realista do Universo dominado por GCG deve porém possuir também um componente bariônico (nem que seja por razões antrópicas), cuja densidade, como veremos, pode ser determinada com boa precisão a partir de observações. Estudaremos a

seguir como os valores dos três parâmetros do GCG e a densidade de bárions determinam a (in)detectabilidade da topologia cósmica em Universos com seções espaciais não planas.

3.4.2 Detectabilidade da topologia em cosmologias de GCG

Enquanto em modelos do tipo Λ CDM a detectabilidade da topologia é determinada por somente dois parâmetros de densidade, em um Universo dominado pelo GCG esta é determinada por três parâmetros, $\Omega_{\text{Ch}0} = \rho_{\text{Ch}0}/\rho_{\text{crit}}$, A e α (tomaremos um valor fixo para a densidade de bárions, Ω_{b0}). Em vista do que foi feito nas seções anteriores, e para tornar mais fáceis as comparações entre modelos, introduzimos duas novas variáveis,

$$\begin{aligned}\bar{\Omega}_{\Lambda 0} &= \Omega_{\text{Ch}0} A, \\ \bar{\Omega}_{m0} &= \Omega_{\text{Ch}0} (1 - A).\end{aligned}\tag{3.21}$$

A relação desvio para o vermelho—distância (3.7) então toma a forma

$$\chi_{\text{obs}} = \sqrt{|1 - \Omega_0|} \int_0^{z_{\text{obs}}} \left\{ \Omega_{\text{Ch}} + \Omega_{b0}(1+z)^3 + (1 - \Omega_0)(1+z)^2 \right\}^{-\frac{1}{2}} dz, \tag{3.22}$$

onde $\Omega_{\text{Ch}} = \Omega_{\text{Ch}0} \left[\frac{\bar{\Omega}_{\Lambda 0}}{\bar{\Omega}_{\text{Ch}0}} + \frac{\bar{\Omega}_{m0}}{\bar{\Omega}_{\text{Ch}0}} (1+z)^{3(1+\alpha)} \right]^{\frac{1}{1+\alpha}}$ e $\Omega_0 = \Omega_{\text{Ch}0} + \Omega_{b0}$.

De forma análoga ao que acontece no caso Λ CDM, não é possível integrar analiticamente (3.22) para quaisquer valores dos parâmetros, e em particular para qualquer valor de α . Mas como veremos, χ_{obs} é uma função monotonicamente decrescente de α no intervalo $\alpha \in (-1, \infty)$, e é possível obter expressões analíticas para as curvas de contorno nos casos limite $\alpha = -1$ e $\alpha \rightarrow \infty$ quando $z_{\text{obs}} \rightarrow \infty$. Temos assim uma cota inferior e superior para os valores tomados por χ_{obs} em (3.22) para uma dada escolha dos parâmetros de densidade, e sabemos ainda que para valores intermediários de α , χ_{obs} interpolará monotonicamente entre estes extremos.

Provaremos inicialmente a monotonicidade de χ em função de α . Considere o termo $\Omega_{\text{Ch}} \geq 1$, com $z \geq 0$ e $\bar{\Omega}_{\Lambda 0} \neq 0$. Definindo $u = (1+z)^{3(1+\alpha)} > 1$, é possível calcular diretamente que

$$\frac{\partial^2 \Omega_{\text{Ch}}}{\partial u \partial \alpha} = \frac{\bar{\Omega}_{m0}}{\bar{\Omega}_{\text{Ch}0}} \log \left[\frac{u}{\frac{\bar{\Omega}_{\Lambda}}{\bar{\Omega}_{\text{Ch}0}} + \frac{\bar{\Omega}_m}{\bar{\Omega}_{\text{Ch}0}} u} \right] \geq 0. \tag{3.23}$$

Isto garante que a derivada $\partial \Omega_{\text{Ch}} / \partial \alpha$ é uma função crescente de u . Mas para $u = 1$, é fácil ver que $\partial \Omega_{\text{Ch}} / \partial \alpha = 0$. Assim, $\partial \Omega_{\text{Ch}} / \partial \alpha > 0$ para qualquer $u > 1$, ou $z > 0$, e como consequência Ω_{Ch} é uma função crescente de α . Como Ω_{Ch} é o único termo em (3.22) que depende de α , claramente temos que χ_{obs} é uma função decrescente de α . De forma sucinta, para valores fixos de $\bar{\Omega}_m$ e $\bar{\Omega}_{\Lambda}$ a detectabilidade de uma dada topologia se torna menos provável à medida que α aumenta.

Calculamos agora as fórmulas analíticas para as curvas de contorno de (3.22) nos extremos $\alpha \rightarrow \infty$ e $\alpha = -1$. Como vimos na seção anterior, para o caso $\alpha = 0$ as curvas de contorno são bem aproximadas pela reta secante dada por (3.12) - (3.14).

No limite $\alpha \rightarrow \infty$ temos que $\Omega_{\text{Ch}} \rightarrow \Omega_{\text{Ch}0}(1+z)^3$, e portanto χ_{obs} se torna uma função simples de Ω_0 . A integral é funcionalmente idêntica àquela obtida no caso ΛCDM com $\Omega_{m0} = 0$, cujo resultado é dado por (3.10). Tomando o limite $z_{\text{obs}} \rightarrow \infty$ e resolvendo então a equação $\chi_{\text{obs}}(\Omega_0) = r_{\text{inj}}^M$ para Ω_0 , verificamos que as curvas de contorno para $\alpha \rightarrow \infty$ são retas

$$\begin{aligned} \bar{\Omega}_{m0} + \bar{\Omega}_{\Lambda0} &= \text{sech}^2\left(\frac{r_{\text{inj}}^M}{2}\right) - \Omega_{b0} , & \text{para } \Omega_0 < 1 , \\ \bar{\Omega}_{m0} + \bar{\Omega}_{\Lambda0} &= \sec^2\left(\frac{r_{\text{inj}}^M}{2}\right) - \Omega_{b0} , & \text{para } \Omega_0 > 1 . \end{aligned} \quad (3.24)$$

$(\alpha \rightarrow \infty)$

Note que, devido a monotonicidade, (3.24) define uma condição suficiente para indetetabilidade da topologia, análoga à (3.12).

O cálculo do limite $\alpha \rightarrow -1$ é um pouco mais trabalhoso. Tomando o limite $\alpha \rightarrow -1$ na expressão de Ω_{Ch} , e lembrando a definição usual do número de Euler, e , obtemos

$$\Omega_{\text{Ch}} = (1+x)^{3\frac{\bar{\Omega}_{m0}}{\bar{\Omega}_{\text{Ch}0}}} ; \quad (3.25)$$

substituindo a equação acima em (3.22), podemos com algum esforço integrar e resolver a equação resultante para $\Omega_{\text{Ch}0}$, e obter finalmente

$$\begin{aligned} \bar{\Omega}_{m0} + \bar{\Omega}_{\Lambda0} &= \sec^2(C r_{\text{inj}}) + \tan^2(C r_{\text{inj}}) \bar{\Omega}_{\Lambda0} , & \text{para } \Omega_{\text{Ch}0} < 1 \text{ e } C > 0 , \\ \bar{\Omega}_{m0} + \bar{\Omega}_{\Lambda0} &= \text{sech}^2(C r_{\text{inj}}) - \tanh^2(C r_{\text{inj}}) \bar{\Omega}_{\Lambda0} , & \text{para } \Omega_{\text{Ch}0} > 1 \text{ e } C > 0 , \\ \bar{\Omega}_{m0} + \bar{\Omega}_{\Lambda0} &= 1 , & \text{para } C \leq 0 , \end{aligned} \quad (3.26)$$

$(\alpha \rightarrow -1)$

onde $C = (3\bar{\Omega}_{m0}/2\Omega_{\text{Ch}0} - 1)$. A curva de contorno definida por (3.26) coincide portanto com a reta secante do caso ΛCDM quando $\bar{\Omega}_{m0} = 0$, e com a reta plana para $\bar{\Omega}_{m0} \geq \bar{\Omega}_{\Lambda0}$. Na ausência de matéria bariônica, os valores compatíveis com as observações disponíveis implicam em $C \leq 0$, correspondente a $A \geq 2/3$, e neste caso *qualquer* topologia de um espaço curvo seria detectável em princípio. Mas como veremos, mesmo um pequeno valor de Ω_{b0} torna várias topologias indetectáveis. A consequência, um tanto surpreendente, é

que o valor exato da densidade bariônica é um fator crucial para determinar a detectabilidade da topologia cósmica em Universos dominados pelo GCG.

A Figura 3.5 mostra esquematicamente a relação entre detectabilidade e as curvas de contorno no plano paramétrico $\bar{\Omega}_{m0} - \bar{\Omega}_{\Lambda 0}$. As curvas de contorno associadas aos valores $\alpha = -1, 0$ e ∞ , assim como a reta secante, também são indicadas. Para qualquer destes casos, se $\bar{\Omega}_{m0}$ e $\bar{\Omega}_{\Lambda 0}$ assumirem valores entre a curva de contorno respectiva e a reta plana, então a topologia de M é indetectável.

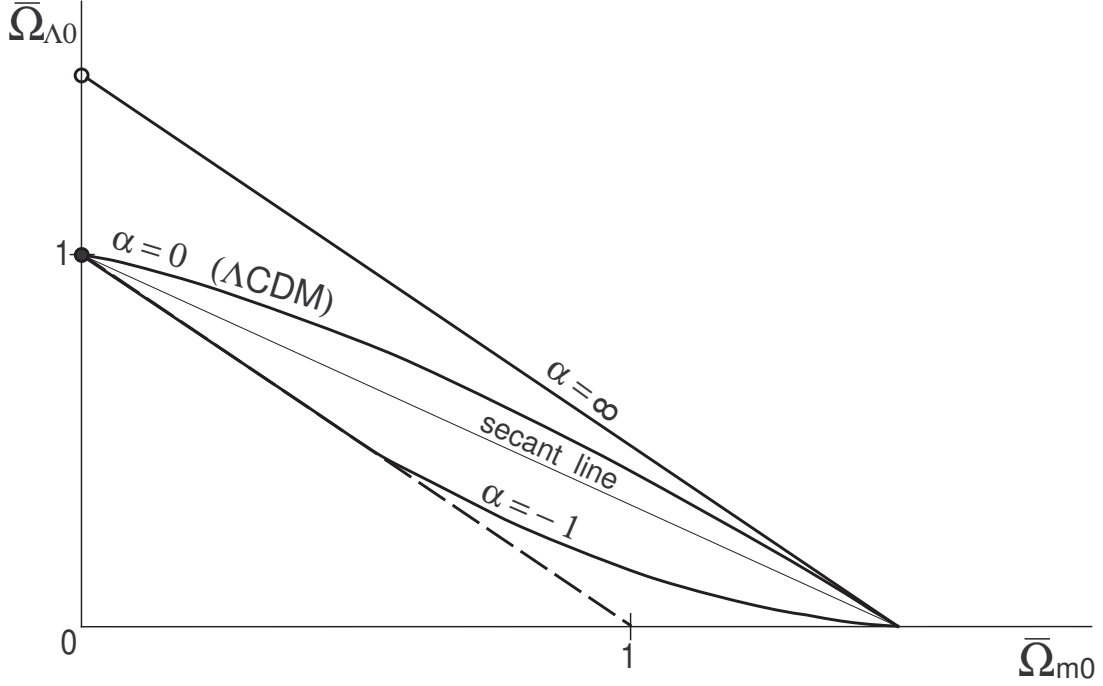


Figura 3.5: Representação esquemática das curvas de contorno $\chi_{obs} = r_{inj}^M$ para $\alpha = -1, 0$ e ∞ , com $z_{obs} = \infty$ e $\Omega_{b0} = 0$. Curvas associadas a valores intermediários de α interpolarão de forma monotônica entre estes extremos. O uso de $z_{obs} = 1089$ não altera as curvas significativamente. A existência de uma densidade de matéria bariônica impede na prática que a curva para $\alpha = -1$ seja degenerada com a reta plana (tracejada). A reta secante (linha mais fina) é uma aproximação excelente para a curva $\alpha = 0$.

Para valores intermediários de α é preciso computar numericamente χ_{obs} para estudar quantitativamente os efeitos dos valores dos parâmetros do GCG na detectabilidade da topologia. Fixamos $\Omega_b = 0,04$ na equação (3.22), com base nas medições das abundâncias relativas dos elementos leves, em combinação com análise teórica da nucleossíntese primordial que se deu após o Big Bang (veja por exemplo Burles *et al.* [28] e Kirkman *et al.* [85]), juntamente com o valor da constante de Hubble (Freedman [55])³. Para Ω_0

³A determinação da abundância relativa dos elementos leves e a nucleossíntese primordial ([28] e [85]) implicam que $\Omega_{b0}h^2 = 0,0214 \pm 0,0018$, enquanto $h = 0,72 \pm 0,08$ ($h \equiv H_0/100 \text{ km/s/Mpc}$) é o

tomamos $0,99 < \Omega_0 < 1,03$, valores compatíveis com medições da posição do primeiro pico acústico [159] no espectro de potência da RCF (como discutido no Apêndice A), combinado com o estudo da estrutura em grandes escalas obtido por Tegmark no caso Λ CDM [165]. Finalmente, fixamos a densidade total do parâmetro $\bar{\Omega}_{m0}$, de modo que $\bar{\Omega}_{m0} + \Omega_{b0} = 0,3$. Nossos resultados, porém, não são particularmente sensíveis ao valor de $\bar{\Omega}_{m0}$. De fato, de α não assumir valores excessivamente negativos ($\alpha \gtrsim -0,5$), temos variações de ordem $\Delta\chi_{obs}/\chi_{obs} \simeq 0,15 \Delta\bar{\Omega}_{m0}/\bar{\Omega}_{m0}$.

Usando as mesmas variedades analisadas na seção anterior no contexto de modelos Λ CDM, mostramos na Tabela 3.5 as condições para (in)detectabilidade, em termos de Ω_0 , para os valores específicos de α . Tomamos $\alpha = 0, 1$ e -1 , usando as expressões mostradas acima para as curvas de contorno, assim como $\alpha = -1/2$, que corresponde ao melhor ajuste obtido a partir da combinação de diversos observáveis (veja Makler et al. [103] e Zhu [191] para detalhes). Estamos particularmente interessados na influência do valor de α para a detectabilidade, pois o GCG com valores diferentes para este parâmetro grosso modo é protótipo para diversos modelos para o conteúdo de matéria-energia do Universo. Variedades indetectáveis para valores do parâmetro de densidade total compatíveis com observações são mostradas em negrito.

A dependência sensível da detectabilidade da topologia cósmica em relação ao parâmetro α fica evidente na Tabela 3.5. Claramente, modelos diferentes fazendo uso do mesmo conjunto de observações (e.g., Λ CDM e GC) podem ter características bastante diferentes em termos da detectabilidade da topologia. Em particular, valores negativos de α favorecem grandemente a detectabilidade (e neste caso, como vimos, o valor específico de Ω_{b0} também é particularmente crucial). No caso extremo $\alpha = -1$ somente as duas variedades esféricas de maior volume consideradas (que tessalam a 3-esfera em somente quatro cópias do domínio fundamental) permanecem indetectáveis. No caso do Gás de Chaplygin clássico, $\alpha = 1$, a maior parte das variedades hiperbólicas e uma minoria significativa das variedades esféricas consideradas se tornam indetectáveis.

Quantificamos a dependência da detectabilidade em termos de α e Ω_{b0} na Figura 3.6, onde mostramos χ_{obs} como função de α , para alguns valores de Ω_0 . Os gráficos são indicados por linhas sólidas quando $\Omega_{b0} = 0$, e por linhas tracejadas quando $\Omega_{b0} = 0,04$. Note que o valor de χ_{obs} não é muito sensível a α para $\alpha > 0$, e em particular para $\alpha > 1$

 valor medido pelo telescópio espacial Hubble [55]. A combinação dá um valor central $\Omega_{b0} \simeq 0,04$ com incertezas quase desprezíveis. Embora estritamente falando estes valores tenham sido obtidos em um contexto Λ CDM, a recombinação se deu em densidades nas quais o GCG se comporta quase exatamente como matéria escura fria; e a determinação observacional de H_0 é puramente cinemática, e portanto não depende da natureza exata dos componentes de matéria e energia escuras.

Variedades	r_{inj}	$\alpha = -1$	$\alpha = -0.5$	$\alpha = 0$	$\alpha = 1$
$S^3/D_9^*, S^3/Z_{18}$	$\frac{\pi}{18}$	1.001	1.002	1.003	1.004
$S^3/D_6^*, S^3/Z_{12}$	$\frac{\pi}{12}$	1.002	1.004	1.007	1.010
$S^3/I^*, S^3/D_5^*, S^3/Z_{10}$	$\frac{\pi}{10}$	1.003	1.005	1.010	1.015
$S^3/O^*, S^3/D_4^*, S^3/Z_8$	$\frac{\pi}{8}$	1.004	1.008	1.015	1.023
$S^3/T^*, S^3/D_3^*, S^3/Z_6$	$\frac{\pi}{6}$	1.007	1.014	1.026	1.042
$S^3/D_2^*, S^3/Z_6$	$\frac{\pi}{4}$	1.016	1.031	1.059	1.094
m004(1,2)	0.183	0.999	0.998	0.997	0.995
m004(6,1)	0.240	0.998	0.997	0.994	0.991
m003(-4,3)	0.288	0.998	0.996	0.992	0.987
m003(-2,3)	0.289	0.998	0.996	0.992	0.987
m003(-3,1)	0.292	0.998	0.996	0.992	0.987
m009(4,1)	0.387	0.996	0.992	0.985	0.977
m007(3,1)	0.416	0.995	0.991	0.983	0.974

Tabela 3.5: Valores mínimos (máximos) para Ω_0 que garantem detectabilidade em princípio para variedades esféricas (hiperbólicas) selecionadas em modelos GCG. Tomamos $\bar{\Omega}_M = 0,26$, $\Omega_b = 0,04$ e $z = 1089$. Números em negrito indicam que as variedades correspondentes são indetectáveis para $0,99 < \Omega_0 < 1,03$.

($\Delta\chi_{obs}/\chi_{obs} \simeq 0,3$ para $1 \leq \alpha \leq \infty$). Isto implica que as condições de indetectabilidade (3.24), obtidas para $\alpha \rightarrow \infty$, não são muito mais restritivas do que aquelas derivadas para um Universo dominado pelo gás de Chaplygin clássico. Neste caso a presença de um pequeno componente bariônico Ω_{b0} tampouco altera substancialmente χ_{obs} . Para $\alpha < 0$, por outro lado, χ_{obs} é extremamente sensível a ambos os parâmetros α e Ω_{b0} .

3.5 Considerações gerais sobre a detectabilidade

Vimos que as condições de detectabilidade de uma possível topologia não trivial das seções espaciais do Universo dependem fundamentalmente do modelo cosmológico escolhido. Além disso, para serem de utilidade prática, tais condições precisam levar em conta as

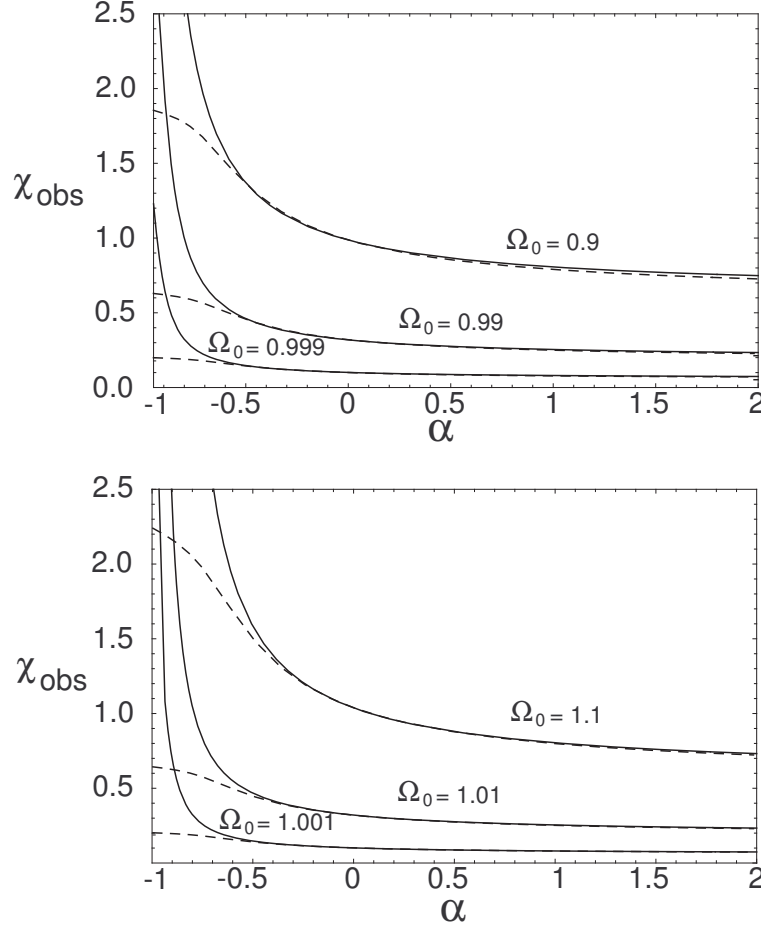


Figura 3.6: O valor de χ_{obs} para diversos valores de $\Omega_0 = \Omega_{Ch0} + \Omega_{b0}$ em Universos localmente hiperbólicos (acima) e esféricos (abaixo). Tomamos $z_{obs} = 1089$ e $\bar{\Omega}_{m0} + \Omega_{b0} = 0, 3$. O componente bariônico é desprezado nas linhas sólidas e levado em conta nas linhas tracejadas.

incertezas observacionais associadas aos parâmetros deste modelo. O presente capítulo mostrou como fazer isto sistematicamente. Estudamos a detectabilidade da topologia das seções espaciais em duas classes de modelos de grande interesse, com base nas curvas de contorno da relação desvio para o vermelho—distância no espaço paramétrico. Ficou bastante claro que uma mesma topologia pode ter condições de detectabilidade diferentes para modelos cosmológicos distintos.

O texto acima inclui alguns exemplos de aplicações dos métodos apresentados. Uma outra aplicação pode ser encontrada no Capítulo 5. Mas, fundamentalmente, estes métodos são ferramentas simples, a serem usados para deduzir quais topologias podem ser excluídas por observações, levando em conta a escolha de modelo e erros observacionais; e quais topologias são indecidíveis, ou seja, não passíveis de detecção, e portanto de serem excluídas observacionalmente. Esta abordagem pode ser facilmente estendida para estudar a detectabilidade em qualquer outro modelo cosmológico cuja relação desvio para o

vermelho—distância possa ser parametrizada em um número razoável de parâmetros. Temos então condições de determinar, em termos das incertezas nos valores dos parâmetros do modelo, sob quais condições uma dada topologia pode ser detectável ou não.

Podemos ainda formular o problema inverso da detectabilidade da cosmologia: Se a topologia (não trivial) do Universo for determinada, quais restrições serão impostas aos parâmetros de um dado modelo cosmológico? É para esta questão que nos voltamos no capítulo seguinte.

Capítulo 4

O problema inverso em topologia cósmica

Nemo enim fere saltat sobrius nisi forte insanit

(Geralmente ninguém dança sóbrio, a não ser que por acaso tenha enlouquecido.)

Marco Túlio Cícero

4.1 Introdução

No capítulo anterior, vimos como a escolha do modelo para o conteúdo material do Universo, o valor de seus parâmetros e as incertezas observacionais associadas a sua determinação, determinam sob que condições uma possível topologia não trivial para as seções espaciais é detectável ou não. Topologias detectáveis podem, em princípio, ser excluídas por observações. Como vimos no Capítulo 3, os valores dos parâmetros observacionais podem ser usados para separar as topologias possíveis dentre aquelas já excluídas por observação, as possíveis de serem detectadas e as indetectáveis. Da mesma maneira, uma eventual detecção da topologia pode ser usada para se restringir parâmetros cosmológicos. Este é o chamado problema inverso da topologia cósmica.

A determinação da topologia cósmica, além de uma importante e fundamental descoberta científica por si só, também seria uma medida observacional, assim como foram a determinação da constante de Hubble ou do espectro angular de potência da radiação cósmica de fundo. E, da mesma maneira que estas medidas, poderia ser usada para definir

com maior precisão o Universo em que vivemos. Em seu nível mais básico, a topologia define univocamente o espaço de cobertura, e portanto o sinal da curvatura espacial. Determinar o sinal da curvatura espacial já seria um resultado significativo, já que um Universo quase plano (consequência típica de modelos inflacionários) provavelmente torna a questão do sinal da curvatura indecidível. Em particular, no espaço dos parâmetros de densidade, um Universo exatamente plano tem medida nula, e portanto sua existência não poderia jamais ser comprovada medindo-se somente tais parâmetros. Mas a detecção de um número suficiente de imagens múltiplas pode determinar univocamente o grupo de holonomias, caracterizando além de qualquer incerteza a planitude do espaço. Por sua vez o uso da planitude como prior¹ (e não somente hipótese de trabalho) restringe [158] enormemente as regiões de confiança observacionais. Assim, a determinação de uma topologia plana seria um grande triunfo para a cosmologia de precisão, com um impacto talvez comparável à observação das anisotropias da radiação cósmica de fundo.

Se soubéssemos exatamente qual é a topologia das seções espaciais do Universo, seríamos capazes de determinar o sinal da curvatura espacial, que é determinado pelo espaço de cobertura ($\mathbb{R}$, $\mathbb{S}$ ou $\mathbb{H}$, c.f. Apêndice B) da variedade topológica em questão. O valor de Ω_0 ficaria então restrito, seja a um dos semi-planos $\Omega_0 > 0$ ou $\Omega_0 < 0$, seja à curva $\Omega_0 = 0$. Mas como vimos no Capítulo 3, devido à rigidez da topologias não-planas, a sua detecção impõe vínculos significativos ao espaço paramétrico dos modelos cosmológicos. Vínculos estes que podem ser combinados com outras observações para determinar novas e mais precisas regiões de confiança para os parâmetros observacionais. Usando a linguagem do capítulo anterior, se as seções espaciais do Universo são variedades curvas multiplamente conexas, então para um dado modelo cosmológico o valor dos parâmetros possíveis se situa em um semi-plano definido pela curva (ou hiper-superfície quando são mais de 2 parâmetros) de contorno no espaço paramétrico definida pelo raio de injetividade da topologia em questão.

Podemos em alguns casos obter vínculos ainda mais restritivos. Em variedades curvas, devido à sua propriedade de rigidez, as distâncias entre imagens de um mesmo ponto, medidas em unidades do raio de curvatura, são invariantes topológicos conhecidos. Seria portanto possível usar uma ou mais destas distâncias conhecidas entre imagens múltiplas de origem topológica com uma 'régua-padrão' para medir o raio da última superfície de espalhamento, χ_{lss} , em unidades do raio de curvatura. Como vimos no Capítulo 3, fixar o valor de χ_{lss} corresponde a definir uma curva de contorno do tipo $\chi_{lss} = \text{const}$ (ou a uma pequena faixa em torno dela, se levarmos em conta as inevitáveis incertezas observacionais) no espaço dos parâmetros do modelo cosmológico, o que por sua vez

¹Usamos o termo em inglês, pois não há um termos correspondente em português de uso corrente.

determina um vínculo entre estes parâmetros; tal vínculo teria um efeito semelhante sobre as incertezas observacionais ao da hipotética determinação que a curvatura é nula (caso que corresponde à curva de contorno degenerada $\chi_{lss} = 0$).

Mostraremos dois exemplos do uso do problema inverso da detectabilidade da topologia cósmica para produzir vínculos sobre parâmetros cosmológicos, em duas classes de modelos do conteúdo material do Universo que já nos são familiares desde o capítulo anterior. Para o modelo Λ CDM, tomamos como ponto de partida a possível detecção da topologia específica [148] cujo domínio fundamental é o dito dodecaedro de Poincaré [98], para restringir a região permitida no plano paramétrico $\Omega_{\Lambda 0} - \Omega_{m0}$. Em seguida, assumindo um Universo dominado por gás de Chaplygin generalizado (GCG), obteremos restrições sobre os parâmetros do modelo que resultariam da detecção de diversas variedades não-planas. Em ambos os casos, combinamos os vínculos resultantes da detecção da topologia cósmica com dados advindos da observação de supernovas. Estes resultados foram apresentados pela primeira vez em [139] e [101].

Embora atualmente a determinação da topologia cósmica seja somente uma possibilidade, é possível que ela ocorra em breve. Os meios teóricos e observacionais já existem, com ressalvas. Neste caso, a abordagem utilizada seguir seria diretamente aplicável, e os ganhos em precisão reais e muito significativos.

Reafirmamos ainda que os métodos apresentados aqui são totalmente gerais, e poderiam ser aplicados com pequenas modificações para quaisquer outros modelos do conteúdo material do Universo (como foi feito em [13], por exemplo).

4.2 Círculos no céu e supernovas em modelos Λ CDM

Estudaremos as conseqüências observacionais da detecção de uma isometria pertencente ao grupo de holonomias de uma topologia conhecida. Especificamente, trataremos da observação das translações de Clifford obtidas pela identificação das faces do dodecaedro de Poincaré, usando o método de círculos no céu. A escolha da topologia e do método de detecção é motivada por resultados recentes na área, mas não é crucial para a aplicação do método.

Inicialmente, discutiremos brevemente o método, apresentado pela primeira vez em 1998 por Cornish, Spergel e Starkman [39], e o seu uso em algumas buscas preliminares por imagens múltiplas de origem topológica que já foram efetuadas, sem por enquanto obterem resultados positivo.

4.2.1 O método de círculos no céu

A última superfície de espalhamento (USE) se situa à maior distância atualmente acessível a observação². Para além dela (e portanto antes do desacoplamento entre matéria e radiação), o Universo era opaco em todos os comprimentos de onda. Assim, se o Universo possuir seções espaciais com uma topologia não trivial detectável, então haverá algum tipo de repetição de padrão entre diferentes regiões da USE gerada pelo grupo de holonomias. Se tal padrão é na prática observável é uma questão importante, mas a sua existência é uma consequência inescapável de uma topologia não trivial detectável.

Para um observador comóvel O , a USE é uma esfera S , situada em uma seção espacial de tempo cósmico constante definida pelo desvio para o vermelho z_{lss} . Se a variedade que descreve esta seção espacial, $M = \tilde{M}/\Gamma_M$, tem topologia não trivial, então existem holonomias γ e $\gamma^{-1} \in \Gamma_M$, $\gamma^{-1}\gamma = I$, que levam S em cópias de si mesma, γS e $\gamma^{-1}S$. Note que uma topologia detectável corresponde exatamente ao caso no qual S tem interseção não nula com γS . Mas por uma questão de simetria, a interseção de duas esferas em um espaço de curvatura constante é sempre um círculo. Temos assim um par de círculos, $C_1 = S \cap \gamma S$, e $C_2 = S \cap \gamma^{-1}S$, idênticos a menos de fase, pois que são identificados pela isometria, $C_2 = \gamma C_1$. Para o observador comóvel, portanto, cada holonomia detectável implicará a existência de dois círculos de mesmo raio na USE, que apresentam padrões de flutuações de temperatura ao longo de seus perímetros, $\delta T(\theta)$ e $\delta T'(\theta')$, $\theta \in [0, 2\pi)$, idênticos a menos de uma fase, $\delta T(\theta) = \delta T'(\theta' + \phi)$. Veja a Figura 4.1 para uma ilustração.

Considere agora um mapa da radiação cósmica de fundo (RCF), onde não sabemos a priori se o espaço é multiplamente conexo ou não. É possível fazer um levantamento de todos os pares de círculos de mesmo raio, mas para estudá-los é preciso definir algum indicador que discrimine entre pares não relacionados e pares ligados por uma isometria. Cornish, Spergel e Starkman propuseram [39] como indicador uma ligeira modificação da função de correlação de dois pontos restrita a um par de círculos, que denominaram método de 'círculos no céu'. Para círculos de mesmo raio ρ , C_1 e C_2 , centrados em $\hat{c}_1$ e $\hat{c}_2$, e com padrões de temperatura ao longo de seus respectivos perímetros $T_1(\phi_1)$ e $T_2(\phi_2)$ onde $\phi_1, \phi_2 \in [0, 2\pi)$, temos

$$S(\varphi, \rho, \hat{c}_1, \hat{c}_2, \epsilon) = \frac{\langle 2T_1(\epsilon\phi)T_2(\phi + \varphi) \rangle}{\langle T_1(\phi)^2 + T_2(\phi + \varphi)^2 \rangle}, \quad (4.1)$$

onde $\langle \rangle = \int_0^{2\pi} d\phi$, $\epsilon = \pm 1$, e φ é a fase relativa entre os círculos. O indicador $S(\varphi, \epsilon)$ é função da fase relativa φ , dos demais parâmetros que definem o círculo, que determina se

²E provavelmente assim permanecerá, até que as emissões de neutrinos ou ondas gravitacionais primordiais sejam mapeadas em alta resolução.

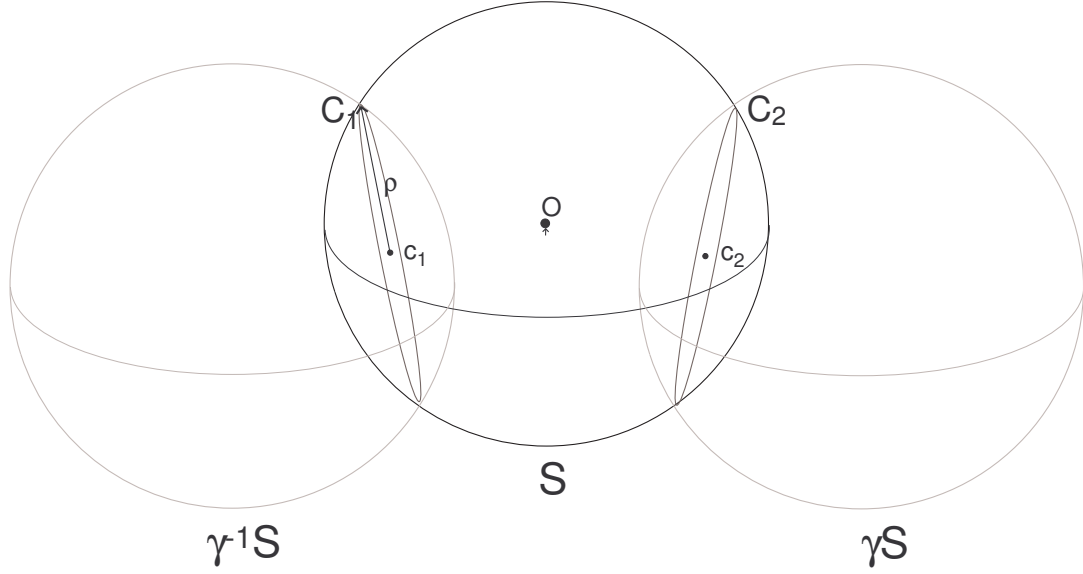


Figura 4.1: Última superfície de espalhamento S em uma 3-variedade $M = \tilde{M}/\Gamma_M$, e suas imagens γS e $\gamma^{-1}S$ pela isometria $\gamma \in \Gamma_M$. As interseções $C_1 = S \cap \gamma S$ e $C_2 = S \cap \gamma^{-1}S$ são círculos, identificados por $C_2 = \gamma C_1$. Note que a posição e fase relativa dos círculos depende da isometria, mas as flutuações de temperatura ao longo de seus respectivos perímetros serão em todos os casos idênticas. Este resultado geométrico se aplica igualmente para $\tilde{M} = \mathbb{R}^3, \mathbb{S}^3$ ou $\mathbb{H}^3$.

a isometria que identifica os círculos preserva orientação ou não (no caso de variedades orientáveis, $\epsilon = 1$).

Em um caso ideal, pares de círculos correlacionados seriam idênticos a menos da fase relativa e orientação. Para tais pares, teríamos portanto $S(\varphi, \epsilon) = 1$ onde ϵ é igual a 1 se a isometria que identifica os círculos preserva orientação, e -1 caso contrário (os demais parâmetros são omitido por simplicidade daqui em diante). Em relação a pares não-correlacionados, se cada círculo consiste de N pixels, $S(\varphi, \epsilon)$ tem média zero e variância $\sim 1/N$. Em [39], Spergel *et al.* calcularam as distribuições para $S(\varphi, \epsilon)$ para pares correlacionados e não-correlacionados, assumindo a existência de um ruído gaussiano, e mostraram que a detecção de pares correlacionados se torna mais provável (e os falsos positivos menos freqüentes) para um maior número de pixels e maior razão sinal/ruído, como seria de se esperar. Concluíram ainda que tais círculos, se existissem, poderiam ser detectados nos mapas de RCF obtidos pelo WMAP.

Na prática, porém, existem alguns complicadores que tornam tal detecção uma questão mais delicada. Em primeiro lugar, podem existir correlações de origem não topológica entre pontos distintos, mesmo em mapas gaussianos. E a própria gaussianidade da RCF já foi posta em dúvida (embora os desvios relatados sejam pequenos e.g. [52]; veja também

a Seção 2.2).

Também foi sugerido que o efeito Doppler advindo do componente radial da velocidade própria da matéria na USE (a velocidade própria em cada ponto terá diferentes projeções radiais para imagens múltiplas vindas de direções diferentes), e o efeito Sachs-Wolfe integrado [110], oriundo da evolução temporal das flutuações de densidade do foreground, poderiam borrar os círculos correlacionados, impedindo a sua detecção [6]. Outros complicadores, tais como o efeito Sunyaev-Zeldovich [190] e a espessura finita da USE, também podem afetar a detecção. Além disso, como vimos no Capítulo 2, a RCF apresenta desvios da isotropia estatística, que além de gerarem mais correlações entre pontos, também sugerem a existência de contaminação de foreground. Finalmente, o movimento próprio de nossa galáxia distorce os círculos no céu, degradando o sinal da detecção, embora, como veremos no Capítulo 6, este último efeito seja totalmente caracterizado, e possa ser corrigido.

As conseqüências de todos estes efeitos para a detectabilidade dos círculos no céu ainda não foram completamente estudadas. As simulações em mapas gerados por Monte Carlo sugerem que o indicador $S(\varphi, \epsilon)$ é robusto, mas por hora a não-detecção de uma topologia não trivial não pode ser tomada como prova definitiva de sua ausência.

Existe ainda um último empecilho de ordem prática ao uso do método de círculos no céu. O espaço paramétrico usado para definir todos os pares de círculos possíveis tem seis dimensões, o que implica que o número de combinações a serem testadas em um mapa de RCF de tamanho razoável é enorme (e aproximadamente proporcional ao cubo do número de pixels utilizado). Computacionalmente, calcular $S(\varphi, \epsilon)$ para todos os pares e fases possíveis em um tempo razoável tem se mostrado difícil sem supor condições adicionais. Mesmo no futuro, o aumento contínuo da capacidade computacional disponível tende a ser mais do que compensado pela disponibilidade de mapas de RCF com resolução cada vez maior [162].

Neste sentido, uma hipótese plausível que restrinja significativamente o volume do espaço paramétrico a ser analisado é de grande valia. Em certos casos (c.f. Capítulo 5), podemos presumir que o espaço é globalmente homogêneo, o que implica em procurar somente por isometrias que são translações de Clifford (tal restrição seria capaz de detectar elementos do subgrupo de isometrias translacionais mesmo em variedades não-homogêneas, como por exemplo os geradores do toro de cobertura em variedades planas). Como pares de pontos ligados por translações de Clifford são sempre separados pela mesma distância, pares de círculos correlacionados serão sempre antipodais.³

³De fato, lembrando a seção 1.4.4, uma translação γ ao longo de um eixo E comuta com uma rotação em torno do mesmo eixo, e portanto uma vizinhança de E suficientemente estreita (na qual podemos

4.2.2 O dodecaedro de Poincaré como candidato

Uma outra maneira de procurar possíveis sinais de uma topologia não trivial do Universo em mapas da RCF é estudar os seus efeitos nos coeficientes da sua expansão em harmônicos esféricos. O domínio fundamental da topologia impõe condições de contorno que definem os modos disponíveis para gerar o espectro das flutuações de densidade no Universo primordial. Quando se dá o desacoplamento entre matéria e radiação, o valor instantâneo destes modos determina os coeficientes, e por consequência o seu espectro de potência angular. Possíveis divergências em relação ao caso simplesmente conexo serão significativos primariamente nos multipolos de ordem mais baixa, que são exatamente os sujeitos a maior incerteza devido à variância cósmica. Mas a análise simultânea de vários multipolos pode resultar em um sinal estatisticamente significativo. Em 2003, J.P. Luminet e colaboradores propuseram que os baixos valores medidos para os coeficientes⁴ dos momentos de quadrupolo ($\ell = 2$) e octopolo ($\ell = 3$) [98] poderiam ser bem explicados se as seções espaciais do Universo fossem o chamado dodecaedro de Poincaré (c.f. o Apêndice B), em que os seis pares de faces opostas pentagonais são identificadas por meio de uma rotação de $\pi/5$, e é uma variedade globalmente homogênea, com um grupo de holonomias I^* gerado por três translações de Clifford independentes. O melhor ajuste foi obtido com $\Omega_0 = 1,013$, consistente com o intervalo dado em [158], $\Omega_0 = 1,015 \pm 0,017$. Como seria de esperar, a proposta criou enorme interesse, e foi estudada em detalhes por vários grupos (veja e.g. [68]). Os estudos mais recentes de Frank Steiner e colaboradores (veja [6] e [7]) indicam que de fato o dodecaedro de Poincaré como poliedro fundamental, assumindo valores fisicamente plausíveis do parâmetro de densidade total, ajusta bem o espectro de potência angular para multipolos de ordem baixa. O melhor ajuste neste caso se dá quando o raio angular dos círculos é $\rho = 50^\circ$. Aurich *et al.* ainda afirmam em [7] que a detecção de círculos no céu é improvável, pois estes seriam borrados pelos efeitos Doppler e Sachs-Wolfe integrado em escalas inferiores a $\delta\rho \simeq 6^\circ$. Cornish e seus colaboradores porém defendem a eficácia do método em [151].

O próximo passo natural foi obviamente procurar detectar diretamente esta topologia. A intersecção do dodecaedro com a USE se dá em seis pares identificados de círculos antipodais de mesmo raio e separação, concêntricos com o centro das faces. Mas uma busca por pares de círculos antipodais e quase antipodais em mapas de RCF gerados sempre incluir um círculo perpendicular C , aumentando o raio da USE) será simétrica por rotações em torno do mesmo. Como o centro do círculo $C_2 = \gamma C$ também deve se situar ao longo de E , a simetria garante que C_2 também será um círculo perpendicular a E .

⁴Foram usados valores obtidos a partir dos dados de 1 ano do WMAP. O coeficiente do hexadecupolo ($\ell = 4$) foi usado como fator de normalização.

a partir do primeiro ano de observações do WMAP não revelou nenhuma correlação significativa [38]. A sensibilidade do método neste caso não permitiria a detecção de círculos com menos do que 18° de raio, porém.

Uma busca subsequente dos círculos [8], usando métodos mais sofisticados de filtragem, e procurando especificamente pelas holonomias dos espaços esféricos binários poliédricos, obteve resultados marginalmente significativos para os padrões gerados tanto pelas holonomias do dodecaedro de Poincaré quanto pelo grupo binário tetraédrico, T^* . Mas os autores não consideram os resultados conclusivos, devido às complexas estruturas de ruído e contaminação de foreground, ainda não completamente entendidos, presentes nos mapas da RCF. Roukema, por outro lado [148], obteve indicações preliminares que parecem sugerir a presença dos círculos, com raio $\rho = 11^\circ \pm 1^\circ$, mas a significância destes círculos foi contestada em [151].

Motivados por este conjunto de resultados, vamos examinar algumas consequências observacionais de um Universo com seções espaciais descritas pelo dodecaedro de Poincaré.

4.2.3 Combinando dados: supernovas tipo SN Ia e a topologia cósmica

Para qualquer observador em um espaço com a forma de um dodecaedro de Poincaré, o comprimento da menor holonomia, $2r_{inj}$ (c.f. a Equação (3.2)), é por simetria simplesmente a distância entre os centros de duas faces opostas do domínio fundamental. Mas, como vimos, estes pontos também são o centro dos círculos no céu que serão detectados. Digamos que tais círculos tenham raio ρ . Eles se situam na interseção da USE, uma esfera com raio χ_{lss} , com os planos onde se situam as faces do domínio fundamental. Os pontos situados no centro do círculo, em um ponto qualquer do seu perímetro, e na posição do observador, determinam um triângulo retângulo esférico, com catetos r_{inj} e ρ , o primeiro dos quais é conhecido e segundo diretamente mensurável, e hipotenusa χ_{lss} . Usando a regra de Napier para tal triângulo, uma das relações fundamentais da trigonometria esférica [40], relacionamos estas quantidades para obter (o problema é ilustrado na Figura 4.2)

$$\cos \rho = \frac{\tan r_{inj}}{\tan \chi_{lss}}, \quad (4.2)$$

onde $r_{inj} = \pi/10$ para o dodecaedro de Poincaré. Para um modelo Λ CDM, χ_{lss} é nossa velha conhecida do capítulo anterior, a equação (3.8), que repetimos por conveniência a seguir,

$$\chi(z_{lss}) = \sqrt{|1 - \Omega_0|} \int_0^{z_{lss}} [(1+x)^3 \Omega_{m0} + \Omega_{\Lambda 0} - (1+x)^2 (\Omega_0 - 1)]^{-\frac{1}{2}} dx. \quad (4.3)$$

onde tomamos novamente $z_{lss} = 1089$.

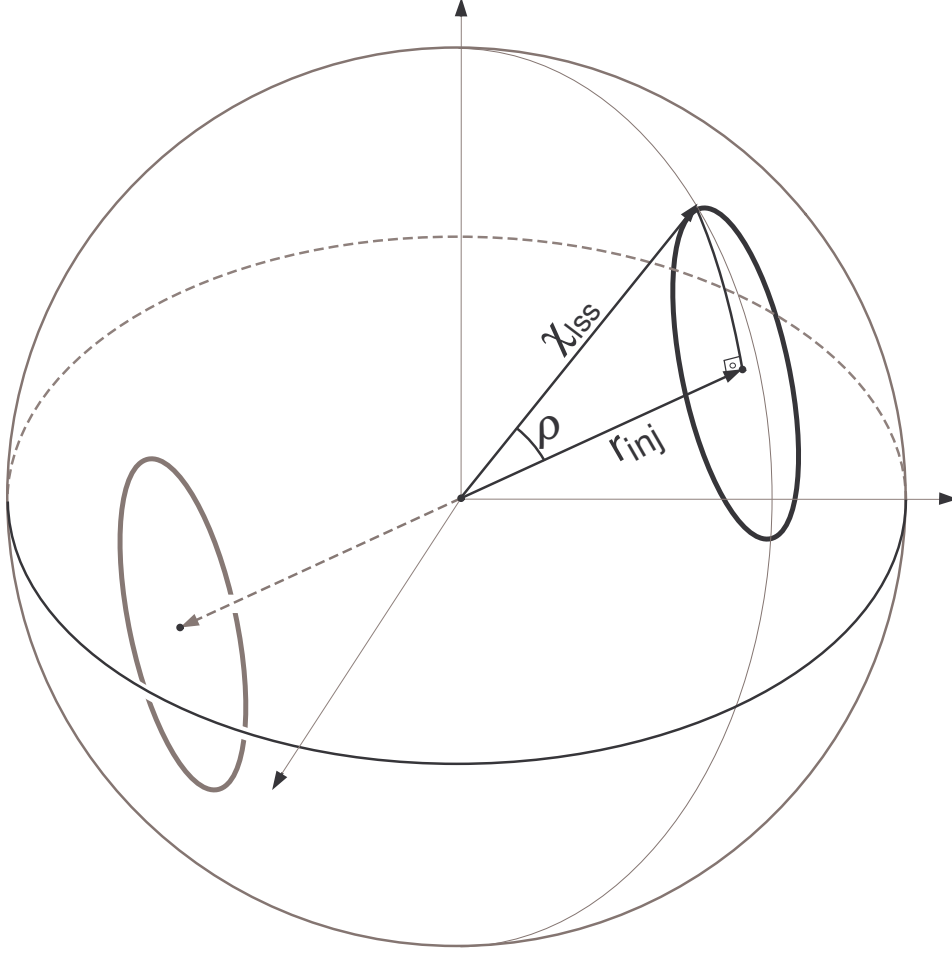


Figura 4.2: Ilustração esquemática de um par de círculos antipodais na última superfície de espalhamento, identificados por uma isometria translacional. A relação entre o raio angular ρ e os comprimentos r_{inj} e χ_{lss} é obtida a partir do chamado pentágono de Napier para triângulos esféricos: $\sin(\pi/2 - \rho) = \tan r_{inj} \tan(\pi/2 - \chi_{lss})$.

As equações (4.2) e (4.3) dão a relação entre o raio angular dos círculos ρ e os parâmetros cosmológicos de densidade Ω_Λ e Ω_m . Podemos então usá-las para restringir estes parâmetros a partir de dados observacionais. Para quantificar esta última afirmativa, é preciso escolher um valor para o raio dos círculos, medido após sua hipotética detecção. Primeiramente, tomamos o valor $\rho_0 = 50^\circ$ obtido por Aurich [6]. Como porém qualquer medição de ρ implica em inevitáveis incertezas observacionais, é preciso ainda quantificar e levar em conta tais erros. Supomos, de forma conservadora, que a incerteza é $\delta\rho \simeq 6^\circ = \sigma_\rho$, a escala abaixo da qual segundo Aurich os círculos seriam borrados.

Na análise estatística a seguir, usamos os dados de supernovas SN Ia obtidas em [144]. Vamos usar especificamente as 157 supernovas que são incluídas na chamada amostra

gold.

No espaço dos parâmetros do modelo Λ CDM, $\Omega_m - \Omega_\Lambda$, as regiões de confiança determinadas pelos dados de supernovas podem ser definidas como curvas de contorno da função de distribuição de probabilidade $\mathcal{L} = \frac{1}{N} \int e^{-\chi_{SN}^2(\Omega_m, \Omega_\Lambda, h)/2} dh$, onde χ_{SN}^2 é a função qui-quadrado dos observáveis referentes às supernovas (no caso, a distância-luminosidade e o desvio para o vermelho). Note que o parâmetro de Hubble $h = H_0 100 \text{ km/s/Mpc}$ é marginalizado (veja o Apêndice A para maiores detalhes). Análises semelhantes foram feitas em estudos recentes, tais como [31], [119] e [3]. O vínculo gerado pela possível detecção dos círculos identificados nas faces do dodecaedro de Poincaré obtido adicionando um termo extra ao qui-quadrado χ_{SN}^2 (c.f. Equação C.15), desta vez usando χ_{lss} como variável.

A Figura 4.3 mostra o resultado da análise conjunta dos dados advindos da medição das SNs Ia e da detecção da topologia cósmica. São mostradas as curvas de contorno de $\mathcal{L}$ correspondente às regiões de confiança para um, dois e três desvios padrão gaussianos (respectivamente 68,3%, 95,4% e 99,7%) no plano paramétrico $\Omega_{m0} - \Omega_{\Lambda0}$. O uso apenas dos dados de SNs Ia, sem nenhuma hipótese cosmológica adicional, é pouco eficiente em restringir o valor de Ω_0 [144]. A introdução de restrições oriundas da topologia cósmica porém reduz significativamente a área das regiões de confiança, e quebra a degenerescência na determinação da densidade total.

A combinação das restrições impostas pelos dados de SNs Ia com a detecção da topologia implica que os parâmetros de densidade que minimizam χ^2 são $\Omega_{m0} = 0,316$ e $\Omega_{\Lambda0} = 0,706$, com $\chi_{min}^2/\nu \simeq 1,13$ (ν é o número de graus de liberdade). Para um nível de confiança de 95,4%, obtemos $\Omega_{m0} = 0,316_{-0,009}^{+0,010}$ e $\Omega_{\Lambda0} = 0,706 \pm 0,010$, e ainda $\Omega_0 = 1,022 \pm 0,014$. Note que o valor de melhor ajuste para a densidade total está em excelente acordo com o obtido a partir das observações do WMAP, $\Omega_0 = 1,015 \pm 0,017$ [158], assim como o valor resultante dos ajustes dos multipolos iniciais do espectro de potência angular do dodecaedro de Poincaré, $\Omega_0 \simeq 1,013$ [98], e $1,015 \leq \Omega_0 \leq 1,020$ [6]. Enfatizamos que nossa análise estatística não faz uso deste último intervalo como prior.

Uma característica notável das regiões de confiança obtidas é que os valores que melhor ajustam tanto Ω_m como Ω_Λ (e por conseguinte Ω_0) têm uma dependência fraca do valor exato do raio medido dos círculos, ρ . Repetindo a análise para os valores sugeridos por Roukema [148], $\rho = 11^\circ \pm 1^\circ$, obtemos $\Omega_{m0} = 0,312_{-0,072}^{+0,078}$, $\Omega_{\Lambda0} = 0,698_{-0,078}^{+0,072}$ e $\Omega_0 = 1,010 \pm 0,002$ para 95,4% de confiança, que são próximos dos valores correspondentes para $\rho = 50^\circ \pm 6^\circ$. Também se nota que o valor da incerteza para o raio determina a largura da região de confiança, mas não muda os valores de melhor ajuste de forma apreciável. Por último, o uso de um vínculo topológico como prior produz um valor para o parâmetro

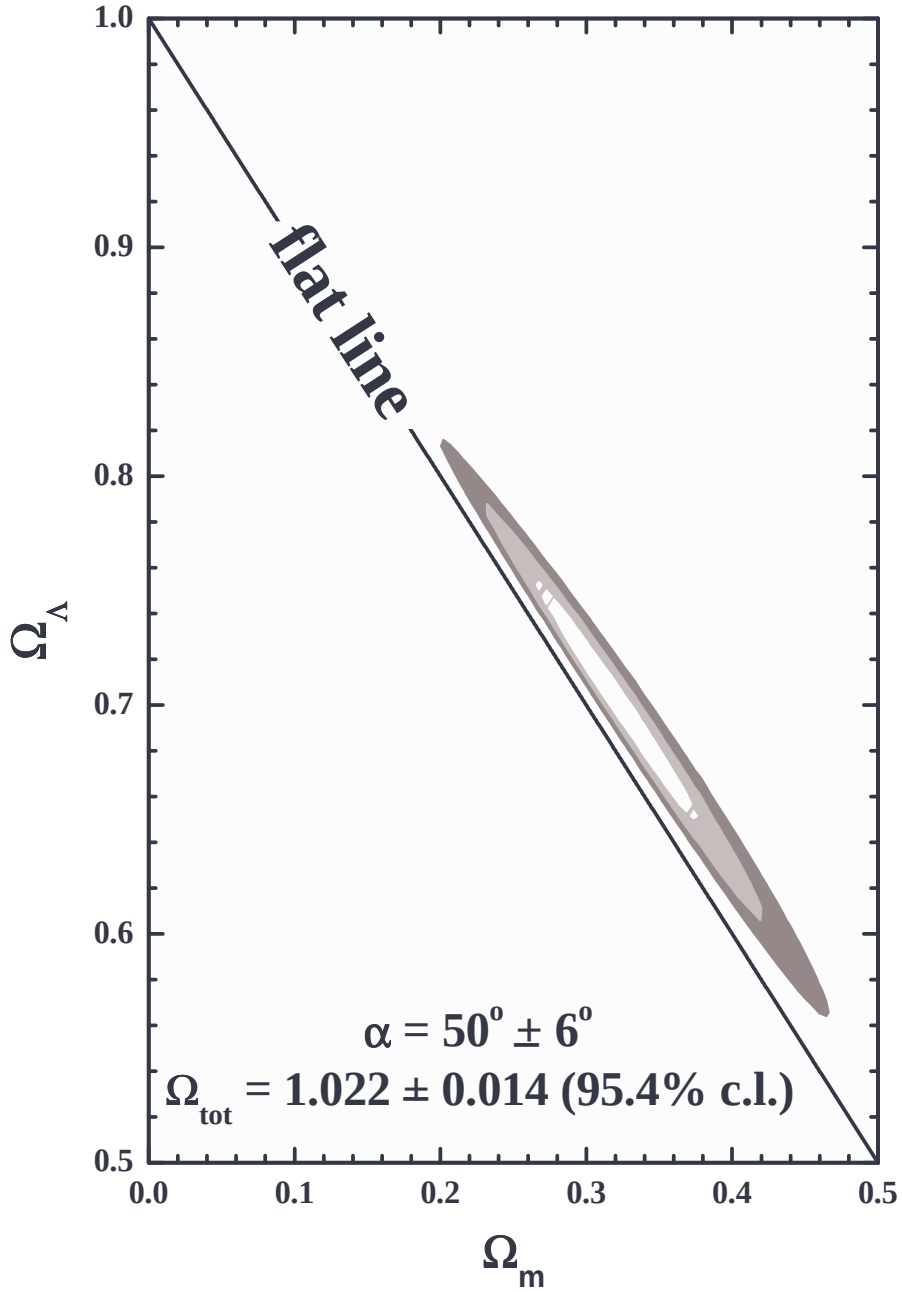


Figura 4.3: Regiões de confiança para 68,3%, 95,4%, e 99,7% no plano paramétrico $\Omega_{m0} - \Omega_{\Lambda 0}$, obtidas combinando medições de supernovas SN Ia e um vínculo resultantes da detecção de uma topologia com seções espaciais dadas pelo dodecaedro de Poincaré. O melhor ajuste se dá para $\Omega_m = 0,316^{+0,011}_{-0,009}$ e $\Omega_\Lambda = 0,706^{+0,010}_{-0,009}$ para um nível de confiança de 95,4% confidence level. O parâmetro de densidade total, assim como os raios angulares dos círculos e a incerteza associada, também são mostrados.

de densidade da matéria Ω_{m0} muito próximo dos obtidos a partir de análises dinâmicas ou estimativas de formação de aglomerados (e.g. [4], [54] e [135]). Isoladamente, a análise de SNs Ia obtém como melhor ajuste para a densidade de matéria $\Omega_{m0} \simeq 0,46$, que diverge cerca de σ do valor central obtido em [54].

4.3 O problema inverso da detectabilidade no gás de Chaplygin generalizado

4.3.1 Escolhendo as variáveis

Como vimos na Seção 3.4.1, um mesmo modelo físico do conteúdo de matéria-energia do Universo pode ser expresso em termos de diferentes conjuntos de parâmetros. Embora muitas das previsões observacionais sejam semelhantes, a escolha de parâmetros pode tornar certas análises mais simples e diretas. No caso do GCG, os parâmetros que usamos anteriormente, $\bar{\Omega}_{m0}$ e $\bar{\Omega}_{\Lambda0}$, têm a virtude da simplicidade, permitindo a obtenção de alguns resultados analíticos em casos limite. No que se segue, vamos mostrar como combinar vínculos topológicos com dados observacionais para impor restrições ao espaço paramétrico de modelos tipo GCG. Faz sentido, portanto, utilizar variáveis que sejam restringidas de forma significativa por diferentes conjuntos de dados observacionais. Como vimos, o parâmetro α é útil para diferenciar entre diferentes classes de modelos, que - em termos da detectabilidade da topologia cósmica - se comportam de formas consideravelmente diferentes. A densidade bariônica, Ω_{b0} , por outro lado é uma grandeza física que pode ser medida com alguma precisão (c.f. nota na página 88). Uma escolha natural, portanto, caso obtenhamos vínculos sobre um ou mais parâmetros de densidade, é expressar a densidade total de GCG, Ω_{Ch} , em termos destes últimos, de α e de Ω_{b0} .

Tomamos assim como parâmetro $\Omega_{Ch0} = \rho_{Ch0}/\rho_{crit}$, pois sabemos que a medição da posição do primeiro pico acústico restringe essencialmente o parâmetro de densidade total, que no nosso caso é simplesmente $\Omega_0 = \Omega_{Ch0} + \Omega_{b0}$. Por outro lado, observações da estrutura da matéria em grandes escalas restringem principalmente o componente do GCG que se comporta como um fluido sem pressão (veja e.g. [103]), e que portanto evolui proporcionalmente ao cubo do fator de escala, $(a_0/a)^3 = (1+z)^3$. Para encontrar este componente, basta lembrar que ele deve ser dominante em densidades mais altas, correspondentes a altos valores de z , mas após a recombinação. No limite de alto z , (3.20) se torna

$$\Omega_{Ch}|_{z \rightarrow \infty} = \Omega_{Ch0} (1 - A)^{1/(1+\alpha)} = \tilde{\Omega}_{m0} . \quad (4.4)$$

Vemos portanto que $\tilde{\Omega}_{m0}$ é o componente do GCG que se comporta como matéria escura fria⁵. Neste caso, o parâmetro de densidade associado ao gás de Chaplygin na relação desvio para o vermelho—distância (3.22) toma a forma

$$\Omega_{\text{Ch}} = \Omega_{\text{Ch}0} \left[1 - \left(\frac{\tilde{\Omega}_{m0}}{\Omega_{\text{Ch}0}} \right)^{1+\alpha} + \left(\frac{\tilde{\Omega}_{m0}}{\Omega_{\text{Ch}0}} (1+z)^3 \right)^{1+\alpha} \right]^{\frac{1}{1+\alpha}}.$$

Finalmente, definimos também $\Omega_{m0}^{eff} = \bar{\Omega}_{m0} + \Omega_{b0}$, que pode ser entendida como a fração da densidade total que se comporta como matéria sem pressão.

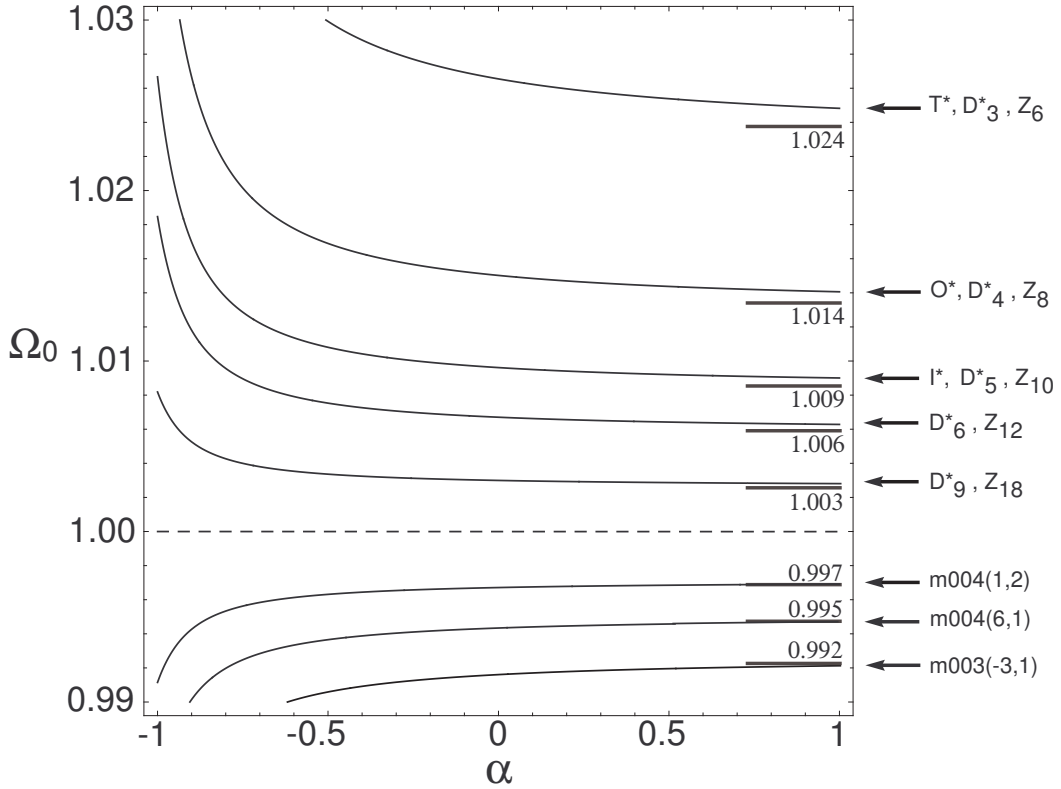


Figura 4.4: Curvas de contorno de $\chi_{obs}(\Omega_0, \alpha) = r_{inj}$ para as variedades esféricas e hiperbólicas citadas no texto. Tomamos aqui $\Omega_m^{eff} = 0,3$, $\Omega_{b0} = 0,04$, e o desvio para o vermelho correspondente à última superfície de espalhamento, $z_{obs} = 1089$.

A Figura (4.4) mostra as curvas de contorno $\chi_{obs} = r_{inj}^M$ no plano paramétrico $\Omega_0 - \alpha$, para Ω_m^{eff} fixo. É possível obter uma desigualdade análoga à (3.23), que garante que χ_{obs} é uma função monotonicamente *crescente*, de α , para $\alpha \in (-1, \infty)$ e valores das densidade

⁵É possível segundo um raciocínio semelhante definir o componente que se comporta como energia escura, que sabemos ser dominante para $z \rightarrow 0$ (i.e., no futuro). Escrevemos então (3.20) neste limite, para obter $\Omega_{\text{Ch}}|_{z \rightarrow 0} = \Omega_{\text{Ch}0} A^{1/(1+\alpha)} = \tilde{\Omega}_{\Lambda 0}$, o componente do GCG que se comporta como constante cosmológica. Mas note que neste caso, $\tilde{\Omega}_{\Lambda 0} + \tilde{\Omega}_{m0} \neq \Omega_{\text{Ch}0}$, então a analogia não é perfeita. Mas no caso $\alpha = 0$, recuperamos $\Omega_{\Lambda 0}$ e Ω_{m0} do modelo Λ CDM.

e z_{obs} fixos. Além disso, é fácil ver que no limite $\alpha \rightarrow 0$ temos simplesmente $\Omega_{Ch} = \Omega_{Ch0}$.

4.3.2 Combinando dados: supernovas tipo SN Ia e a topologia cósmica II

Na seção anterior, vimos como a detecção da topologia impõe restrições na natureza do conteúdo material do Universo. Mas em cosmologia, é comum verificar que somente um observável não é capaz de impor limites significativos para os parâmetros de um modelo. É portanto muito importante combinar as restrições obtidas a partir de múltiplos observáveis, embora este procedimento nem sempre seja trivial. A seguir, vamos combinar as restrições obtidas a partir da observação recente de 194 supernovas do tipo SN Ia (obtidas de [172] e [9]; novamente, veja o Apêndice A para maiores detalhes) com o vínculo topológico $\chi_{obs} > r_{inj}^M$ implícito na possível detecção da topologia M (c.f. o Capítulo 3). Procuramos aqui obter as condições mais gerais possíveis para o componente topológico, e portanto não especificamos outros vínculos explícitos além deste. Assim como fizemos anteriormente, iremos sobrepor as curvas de contorno para algumas variedades não-planas com as regiões de confiança obtidas através de dados de supernovas SN Ia.

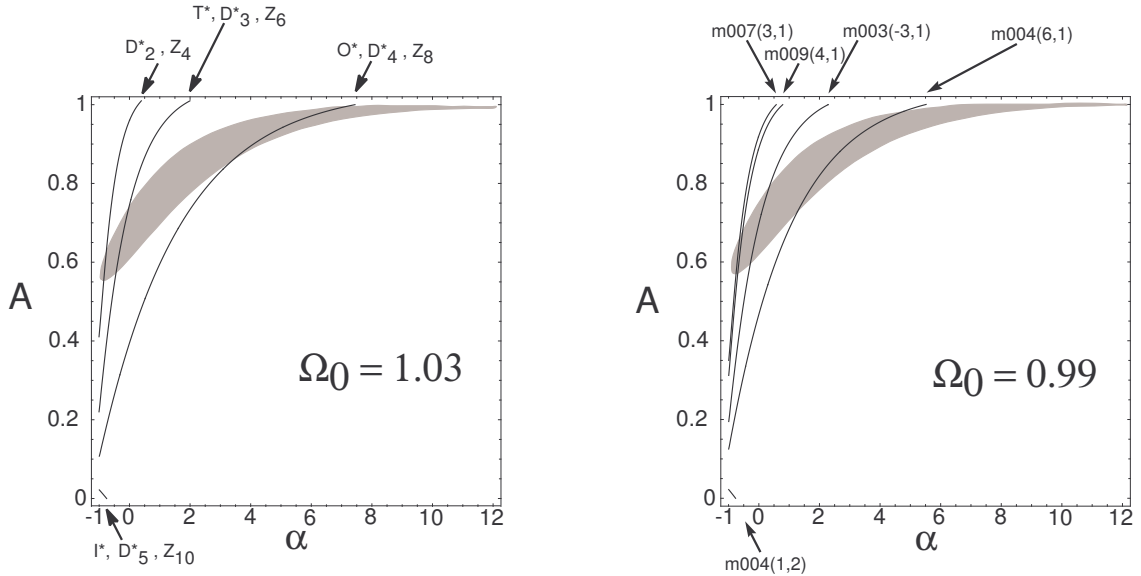


Figura 4.5: Superposição no plano paramétrico $A - \alpha$ da região de confiança a 95% para dados de SNs Ia (cinza claro) e as curvas de contorno $\chi_{obs} = r_{inj}$ para variedades esféricas (esquerda) e hiperbólicas (direita). Tomamos $z_{obs} = 1089$, e fixamos $\Omega_m^{eff} = 0,3$, $\Omega_{b0} = 0,04$.

Na Figura 4.5 mostramos as regiões de confiança em 2σ obtidas a partir de dados de supernovas SN Ia, no plano paramétrico $A - \alpha$, com densidade total fixada em $\Omega_0 = 1,03$ (esquerda) e $\Omega_0 = 0,99$ (direita), e fixando $\Omega_m^{eff} = 0,3$ e $\Omega_{b0} = 0,04$. As formas dos

contornos em ambos os casos são bastante parecidos, pois dados de supernovas não são particularmente sensíveis à densidade total, e podemos supor sem perda significativa de generalidade que estes contornos são válidos para qualquer valor intermediário de Ω_0 . No capítulo anterior, vimos como curvas de contorno $\chi_{obs} = r_{inj}^M$ separam o plano paramétrico $\Omega_{m0} - \Omega_{\Lambda 0}$ em dois semi-planos; um que inclui a reta $\Omega_0 = 1$, no qual a topologia é indetectável, e o seu complemento, onde a topologia é potencialmente detectável. O mesmo acontece no plano paramétrico $A - \alpha$, onde o semi-plano no qual a topologia é não detectável é aquele que inclui o ponto $A = 0, \alpha = -1$. Traçamos neste novo plano paramétrico as curvas de contorno para as mesmas variedades que já havíamos considerado anteriormente, e a região de confiança resultante da observação das SNs Ia. A curva de contorno de cada uma das topologias analisadas também cruza a região de confiança obtida das SNs Ia, separando-a em duas sub-regiões. Na hipótese desta topologia ser detectada, somente uma destas sub-regiões seria admissível. Assim, a combinação de dois conjuntos de observações (no caso, supernovas e topologia) novamente restringe de forma muito mais significativa os parâmetros do modelo GCG do que qualquer um deles isoladamente.

Como pode ser visto claramente na Figura 4.5, quanto maior o valor de r_{inj} , mais significativas se tornam as restrições sobre A e α . Vemos, por exemplo, que com a detecção de $\mathbb{S}^3/D_2^*$ poderíamos descartar o modelo Λ CDM usando apenas dados de supernovas já disponíveis. Além disso, note que a detecção do dodecaedro de Poincaré não impõe restrições adicionais importantes aos parâmetros.

Na Figura 4.6, mostramos as regiões de confiança a 95% de dados de supernovas para os valores de α indicados acima, superpostos às curvas de contorno $\chi_{obs} = r_{inj}$, para algumas variedades não planas selecionadas. Como esperado, os dados de supernovas restringem significativamente A , mas são praticamente indiferentes ao valor de Ω_0 na estreita faixa considerada. Da mesma forma, a detecção de uma topologia não trivial dentre as analisadas por si só imporá somente fracas restrições sobre o valor da densidade total Ω_0 , e nenhuma sobre o valor de A . Porém, o conhecimento da topologia, conjuntamente com os dados de supernovas, restringem ao mesmo tempo A e Ω_0 significativamente.

Para variedades com menor r_{inj} , tais como D_9^* , $L(18, q)$ e $m004(1, 2)$, as restrições ao valor de Ω_0 não variam significativamente com α , como pode ser claramente visto nas figuras. De fato, variedades detectáveis em Universos quase planos não restringem significativamente o parâmetro α . Mas por outro lado, obtemos condições para a densidade total que independem da escolha de α , e são portanto mais gerais. Assim, se as seções espaciais do Universo tiverem sua topologia dada por e.g. $m004(1, 2)$, então forçosamente $\Omega_0 \lesssim 0,998$ com 95% de confiança, para qualquer valor $\alpha > -1/2$.

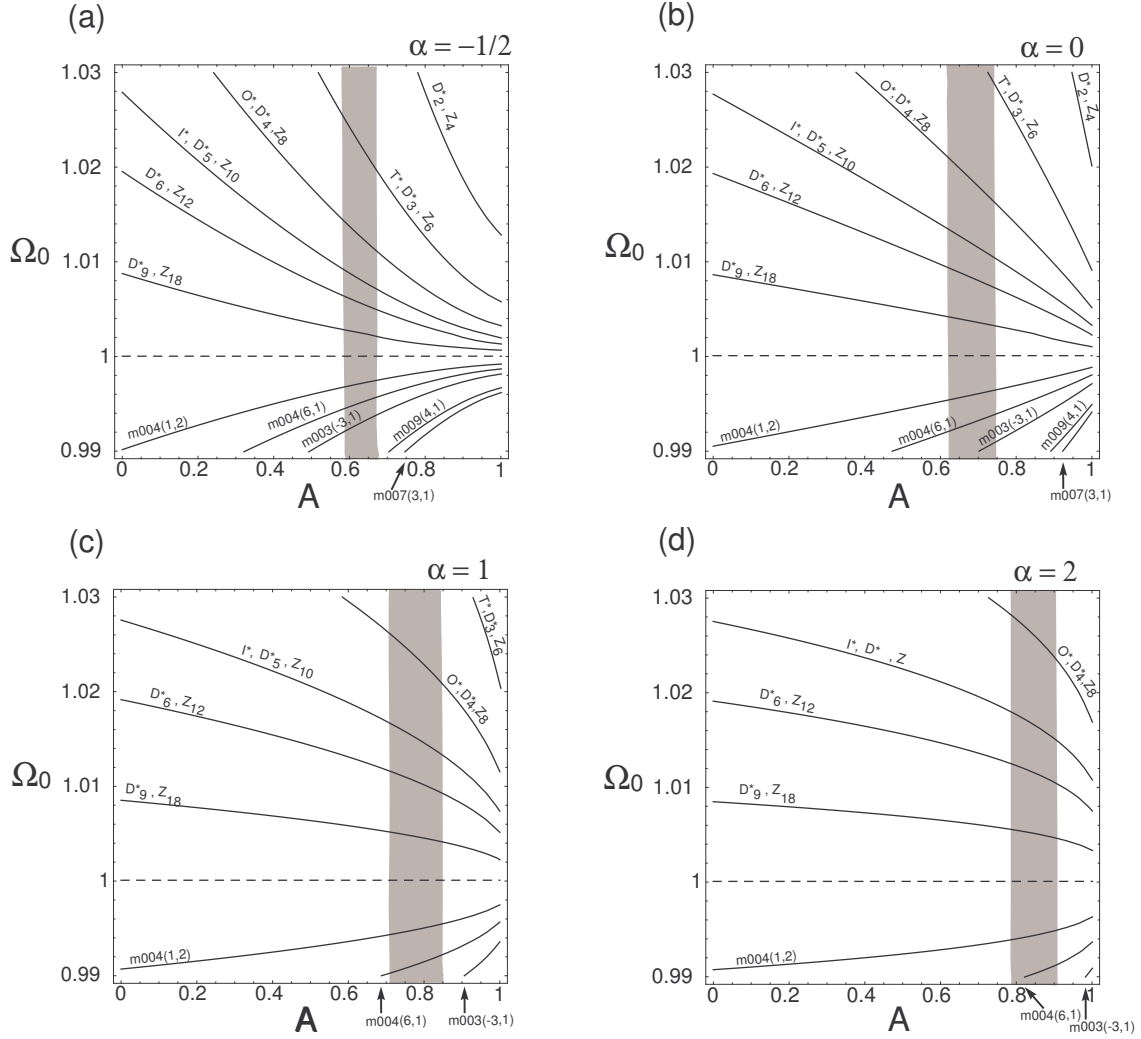


Figura 4.6: Superposição da região de confiança a 95% no plano paramétrico $\Omega_0 - A$ para dados de SNs Ia (cinza claro) e as curvas de contorno $\chi_{obs} = r_{inj}$ para variedades esféricas (esquerda) e hiperbólicas (direita). Novamente, tomamos $z_{obs} = 1089$ e fixamos $\Omega_m^{eff} = 0,3$, $\Omega_{b0} = 0,04$.

Capítulo 5

A forma local do Universo

Deus inventou os inteiros. Todo o resto é trabalho do homem.

Leopold Kroneker

5.1 Introdução

Como já tivemos oportunidade de afirmar, a detecção de imagens múltiplas de origem topológica não determina propriamente a topologia do Universo, mas somente um subconjunto do seu grupo de holonomias (no máximo uma para cada par de imagens detectado), que depende da distância máxima a que têm acesso nossas observações astrofísicas. Para métodos indiretos de detecção, que não dependem da detecção de imagens múltiplas, a distinção entre holonomias detectáveis e não-detectáveis não é tão categórica, mas é possível mostrar que tais métodos não são sensíveis àquelas que geram pares muito além do Universo observável. Isto se torna claro, por exemplo, quando o efeito da topologia no espectro angular de potência da radiação cósmica de fundo é expresso em termos do grupo de holonomias [79]. Outro estudo [142] estimou que holonomias com imagens com separação maior que 10% do diâmetro da última superfície de espalhamento não têm efeito relevante no espectro de potência angular.

O que propomos neste capítulo é caracterizar somente as holonomias que possam ser efetivamente detectadas. Chamaremos este conjunto de 'forma local do Universo'. O seu estudo pode simplificar significativamente o problema da determinação da detectabilidade da topologia cósmica. Como veremos, embora os grupos de holonomia das várias varie-

dades não-planas possíveis sejam bastante diversos, os seus elementos detectáveis muitas vezes têm importantes características em comum, em particular no chamado limite inflacionário, no qual o raio do Universo observável, χ_{obs} , é muito menor que 1 em unidades do raio de curvatura a_0 da seção espacial do Universo. Os resultados a seguir foram apresentados inicialmente em [115], e em seguida em [118], e encontraram interessantes aplicações em e.g. [79]. Alguns resultados novos são apresentados aqui pela primeira vez.

Para cada ponto em um espaço multiplamente conexo, podemos classificar as holonomias em ordem crescente da separação entre o ponto de origem e sua imagem topológica associada (ou, posto de outra maneira, por ordem de comprimento das geodésicas fechadas). Para que a topologia seja detectável, como especificamos no Capítulo 3, é necessário que pelo menos a menor destas distâncias seja menor que o diâmetro do Universo observável (uma esfera de raio χ_{obs} , que, salvo afirmativa em contrário, tomaremos como a distância até a última superfície de espalhamento). De maneira mais geral, a forma local do Universo consiste naquelas holonomias que geram localmente pares de imagens com separação menor do que $2\chi_{obs}$.

É difícil fazer afirmativas categóricas sobre a forma local de variedades planas, pois estas não possuem uma escala característica com a qual o comprimento das geodésicas fechadas possa ser comparado¹ (c.f. o Apêndice B para uma discussão da classificação das formas espaciais para diferentes sinais da curvatura).

Em espaços de curvatura negativa, não há ainda uma classificação geral dos grupos de holonomias possíveis, e pouco pode ser dito sobre elas de aplicação universal. Além disso, eles não admitem translações de Clifford e, portanto, mesmo para cada variedade em particular não é possível calcular vínculos sobre os comprimentos das geodésicas fechadas que se apliquem ao espaço todo.

Variedades esféricas, por outro lado, têm uma classificação completa bem entendida. De fato, o espaço de cobertura é a 3-esfera $\mathbb{S}^3$, cujo grupo de simetrias é isomorfo à $SO(4)$. Os subgrupos finitos e sem pontos fixos deste grupo, que correspondem cada um a uma forma espacial esférica, são conhecidos há quase um século [167], e o fato do espaço de cobertura ser compacto garante que eles podem sempre ser decompostos em subgrupos cíclicos. Podemos sempre expressar uma isometria esférica como uma aplicação cíclica sobre um par de números complexos; uma variedade cujo grupo de holonomia é cíclico, gerado por uma tal isometria, é chamado espaço-lente. Resultados obtidos para tais espaços têm portanto aplicação mais geral.

¹Um 3-toro, por exemplo, possui como geradores três translações com direções e comprimentos totalmente arbitrários (exceto que obviamente são linearmente independentes). Outras variedades planas têm vínculos entre os geradores, mas há sempre pelo menos um parâmetro livre.

Mostraremos como obter a holonomia que gera a geodésica fechada de menor comprimento em cada ponto de um espaço-lente. Apresentaremos também uma caracterização completa dos subgrupos translacionais dos grupos de holonomia esféricos.

No limite inflacionário, a forma local do Universo para variedades esféricas, assim como para as hiperbólicas, se simplifica sensivelmente. Veremos que, quando o módulo da curvatura espacial é suficientemente pequeno, as holonomias detectáveis serão indistinguíveis de translações planas em quase todo o espaço, e terão no máximo dois geradores (i.e., a forma local do Universo será indistinguível de $\mathcal{T}_1 = T^2 \times \mathbb{R}^1$ ou $\mathcal{J}_0 = \mathbb{S}^1 \times \mathbb{R}^2$). Este resultado simplifica deveras a busca por confirmação observacional da topologia, desde que sejam válidos os modelos de inflação genéricos. Relaxando um pouco as restrições sobre a curvatura, podemos aproximar isometrias não-planas detectáveis por isometrias planas do tipo helicoidal. Mostraremos ainda que espaços-lente genéricos, ao contrário das variedades hiperbólicas, sempre possuem geodésicas fechadas 'curtas', com uma cota superior proporcional ao inverso da raiz quadrada da ordem do grupo. Isto implica que, se a topologia é detectável para alguma observador na variedade, ela em geral também o será para todos os observadores, o que é exatamente o oposto do que acontece em variedades hiperbólicas.

Muitos destes resultados são obtidos por meio da teoria dos números, uma área da matemática pura que não é famosa por sua aplicabilidade à cosmologia. São resultados relativamente simples, mas de aplicação bastante geral. Para preservar a fluidez do texto, as provas, assim como uma breve revisão dos elementos relevantes da teoria dos números, são radicadas no Apêndice D.

5.2 A forma local em espaços de curvatura positiva

5.2.1 Caracterizando as holonomias translacionais

Uma isometria da 3-esfera sem pontos fixos pode ser expressa, sem perda de generalidade, como uma ação cíclica de um operador $\gamma_{(p,q)}$ sobre um par ordenado de complexos $z = (z_1, z_2) \in \mathbb{C}^2$ que representam um ponto em $\mathbb{S}^3$ ($|z_1|^2 + |z_2|^2 = 1$), i.e.

$$\gamma_{(p,q)}^n(z_1, z_2) = (e^{2\pi i \frac{n}{p}} z_1, e^{2\pi i \frac{nq}{p}} z_2) ; z_1, z_2 \in \mathbb{C}, |z_1|^2 + |z_2|^2 = 1, \quad (5.1)$$

onde os inteiros $p > 1$ e $0 < |q| < p/2$ não tem fatores em comum (i.e., $(p, q) = 1$) e $\gamma_{(p,q)}^n$ representa a n -ésima aplicação de $\gamma_{(p,q)}$ sobre o par ordenado (z_1, z_2) . A troca do sinal

de q produz isometrias que são imagens espelhadas uma da outra. Todo grupo cíclico² de holonomias esféricas pode ser expresso em termos da ação de um gerador $\gamma_{(p,q)}$ (i.e., com $n = 1$ em (5.1)) para alguma escolha de p e q (veja o Lema D.6 no Apêndice D e a discussão imediatamente anterior).

A distância entre um ponto z e sua imagem por $\gamma_{(p,q)}^n$ é o comprimento de arco ao longo de $\mathbb{S}^3$ (parametrizada como uma 3-esfera imersa em $\mathbb{R}^4$) subtendido pela diferença entre os vetores z e $\gamma_{(p,q)}^n z$, e pode ser expresso como

$$\cos(d(z, \gamma_{(p,q)}^n z)) = \cos\left(\frac{2n\pi}{p}\right) |z_1|^2 + \cos\left(\frac{2\pi nq}{p}\right) |z_2|^2. \quad (5.2)$$

Lembramos que $\cos x$ é uma função estritamente decrescente de x para $0 < x < \pi$, e π é a distância máxima entre pontos na 3-esfera. Assim, a distância entre um ponto e sua imagem topológica pela isometria $\gamma_{(p,q)}$ é uma função decrescente em p e crescente em $|q| \bmod p$. Este valor também depende do módulo $|z_1|$ (ou, de forma equivalente, $|z_2|$), exceto quando $q = \pm 1$. É fácil ver portanto que $\gamma_{(p,q)}$ é uma translação de Clifford se e somente se $q = \pm 1$; neste caso, é arbitrária a escolha do par de círculos máximos da 3-esfera definidos por $z_1 = 0$ e $z_2 = 0$ (os ditos 'equadores', em torno dos quais a ação da isometria se assemelha a uma hélice, girando ora para a esquerda, ora para a direita, dependendo do sinal de q). Uma forma espacial cujo grupo de holonomias seja gerado pela aplicação sucessiva de uma isometria do tipo (5.1) é conhecido como espaço-lente, ou $L(p, q)$; mas em geral o grupo de holonomias de qualquer forma espacial esférica pode ser obtido a partir de um conjunto finito de tais geradores (não necessariamente com os mesmos equadores). Em geral, cada subgrupo cíclico pode ser obtido de forma equivalente a partir de dois geradores diferentes na forma (5.1): O próprio $\gamma_{(p,q)}$, e o seu conjugado $\gamma_{(p,q)}^{q'}$, onde $q'q \equiv 1 \bmod p$. É possível mostrar (c.f o Teorema de Euler D.1 no Apêndice D) que $q' = q^{\phi(p)-1}$, onde $\phi(p)$ é a cardinalidade do conjunto de números inteiros menores que e sem fatores comuns com p . O gerador conjugado do subgrupo cíclico tem então a forma

$$\gamma_{(p,q)}^{\phi(p)-1}(z_1, z_2) = \gamma_{(p,q')}(z_2, z_1) = (e^{2\pi i \frac{1}{p}} z_2, e^{2\pi i \frac{q'}{p}} z_1). \quad (5.3)$$

Para maiores detalhes, veja os Apêndices D e B.

Em variedades planas, como vimos no Capítulo 1, um conceito muito útil é o chamado 'toro de cobertura'. Para qualquer forma espacial com grupo de holonomia Γ agindo

²Um grupo gerado por uma única isometria γ tal que $\gamma^n = I$ para algum $n \in \mathbb{N}$ é chamado grupo cíclico. Como o espaço de cobertura $\mathbb{S}^3$ é compacto, então o grupo gerado por qualquer $\gamma_{(p,q)}$ é de ordem finita, e é portanto cíclico.

sobre $\mathbb{R}^3$, existe sempre um subgrupo composto de translações, Γ_T , tal que o quociente $\mathbb{R}^3/\Gamma_T$ é um toro. Mesmo quando isometrias não-translacionais forem detectáveis, a sua detectabilidade depende da posição do observador na variedade, que obviamente não é conhecida a priori. Mas o toro de cobertura tem sempre a mesma forma para todos os observadores.

Em variedades não globalmente homogêneas, que são exatamente aquelas que têm isometrias não-translacionais no seu grupo de holonomias, a forma local depende da posição do observador (ou, mais especificamente, do módulo $|z_1|$). Seria portanto bastante útil definir o conceito análogo ao toro de cobertura para formas espaciais esféricas. Isto é exatamente o que faremos a seguir. Podemos provar (veja D.9) que para qualquer grupo cíclico de holonomias gerado por $\gamma_{(p,q)}$ (e por consequência para qualquer espaço-lente $L(p, q)$) que

Lema 5.1 *Um espaço-lente $L(p, q)$ possui translações de Clifford como holonomias se e somente se $r(q^2 + 1) \equiv 0 \pmod{p}$ para algum $r < p$. Os subgrupos translacionais terão geradores γ^n e/ou γ^m , onde $n = p/(p, q - 1)$ e/ou $m = p/(p, q + 1)$, com $n, m < p$.*

Note que pelo lema acima serão, no máximo, dois os subgrupos das translações de Clifford por grupo cíclico, com geradores $\gamma_{(p,q)}^n$ e $\gamma_{(p,q)}^m$ que geram respectivamente todas as translações à direita e à esquerda. O equivalente ao toro de cobertura é a composição dos domínios fundamentais de espaços-lente de ordem p/n e p/m centrados em z . Nos equadores, o eixo de ambos coincide, e a interseção se reduz ao espaço-lente de maior ordem (i.e., p/n , se $n \leq m$), enquanto em $|z_1| = |z_2|$ os eixos são ortogonais. A principal diferença neste caso em relação ao toro de cobertura plano é que nem toda forma espacial esférica admite translações de Clifford como holonomias. No limite inflacionário (veja abaixo), que no caso corresponde a p/n e/ou $p/m \gg 1$, este domínio fundamental se torna indistinguível de $\mathcal{J}_0$ ou $\mathcal{T}_1$ (c.f. o Apêndice B para as definições), cujo grupo de holonomias é gerado por uma translação ou por duas translações perfazendo um ângulo de $\frac{1}{2} \arccos |z_1|$.

Na Tabela 5.1 enumeramos todos os espaços-lente com ordem p entre 1 e 24. Mostramos aqueles que são homeomorfos entre si, e também os geradores à esquerda e à direita de suas holonomias translacionais, e indicamos ainda quais espaços-lente são gerados por ação simples, dupla ou composta, segundo a classificação apresentada em [57].

Estes resultados são extensíveis a quaisquer outras formas espaciais esféricas, para as quais holonomias translacionais à esquerda e à direita são geradas pela composição das translações de Clifford indicadas para os subgrupos cíclicos constituintes.

1	2	3	4	5	6
$\mathbb{S}^3$	$L(2, 1) \{ \underline{1} \}$	$L(3, 1) \{ \underline{1} \}$	$L(4, 1) \{ \underline{1} \}$	$L(5, 1) \{ \underline{1} \}$ $L(5, 2)$	$L(6, 1) \{ \underline{1} \}$
7	8	9	10	11	12
$L(7, 1) \{ \underline{1} \}$	$L(8, 1) \{ \underline{1} \}$	$L(9, 1) \{ \underline{1} \}$	$L(10, 1) \{ \underline{1} \}$	$L(11, 1) \{ \underline{1} \}$	$L(12, 1) \{ \underline{1} \}$
$L(7, 2 [3])$	$L(8, 3) \{ -\underline{2} \}$	$L(9, 2 [4]) \{ -3 \}$	$L(10, 3) \{ 5 \}$	$L(11, 2 [5])$ $L(11, 3 [4])$	$L(12, 5) \{ -\underline{2}, 3 \}$
13	14	15	16	17	18
$L(13, 1) \{ \underline{1} \}$	$L(14, 1) \{ \underline{1} \}$	$L(15, 1) \{ \underline{1} \}$	$L(16, 1) \{ \underline{1} \}$	$L(17, 1) \{ \underline{1} \}$	$L(18, 1) \{ \underline{1} \}$
$L(13, 2 [6])$	$L(14, 3 [5]) \{ 7 \}$	$L(15, 2 [7]) \{ 5 \}$	$L(16, 3 [5]) \{ -4 \}$	$L(17, 2 [8])$	$L(18, 5 [7]) \{ -\underline{3} \}$
$L(13, 3 [4])$		$L(15, 4) \{ -3, 5 \}$	$L(16, 7) \{ -\underline{2} \}$	$L(17, 3 [6])$	
$L(13, 5)$				$L(17, 4)$ $L(17, 5 [7])$	
19	20	21	22	23	24
$L(19, 1) \{ \underline{1} \}$	$L(20, 1) \{ \underline{1} \}$	$L(21, 1) \{ \underline{1} \}$	$L(22, 1) \{ \underline{1} \}$	$L(23, 1) \{ \underline{1} \}$	$L(24, 1) \{ \underline{1} \}$
$L(19, 2 [9])$	$L(20, 3 [7]) \{ -5 \}$	$L(21, 2 [10]) \{ -7 \}$	$L(22, 3 [7]) \{ 11 \}$	$L(23, 2 [11])$	$L(24, 5) \{ -4, 6 \}$
$L(19, 3 [6])$	$L(20, 9) \{ -\underline{2}, 5 \}$	$L(21, 4 [5]) \{ 7 \}$	$L(22, 5 [9]) \{ 11 \}$	$L(23, 3 [8])$	$L(24, 7) \{ -\underline{3}, 4 \}$
$L(19, 4 [5])$		$L(21, 8) \{ \underline{3}, -7 \}$		$L(23, 4 [6])$	
$L(19, 7 [8])$				$L(23, 5 [9])$ $L(23, 7 [10])$	

Tabela 5.1: Classificação dos espaços-lente $L(p, q)$, para ordem p entre 1 e 24. Indicamos espaços homeomorfos entre si : Se $L(p, q)$ e $L(p, q')$ são homeomorfos (o que acontece quando $qq' \equiv \pm 1 \pmod{p}$), escrevemos $L(p, q [q'])$. Deixamos implícito que espaços $L(p, q)$ e $L(p, -q)$ são homeomorfos. Para espaços que admitem translações de Clifford à direita e à esquerda como holonomias, indicamos seus geradores entre parênteses, onde o sinal $-$ indica translação à direita. Isometrias que podem gerar a geodésica fechada mais curta em algum ponto são sublinhadas. Variedades com $p = 1$ são geradas por ação simples de quatérnions; variedades geradas por ação dupla estão em negrito. As demais são geradas por ação conjunta. Por exemplo, $L(18, 5 [7]) \{ -\underline{3} \}$ indica que $L(18, \pm 5)$ e $L(18, \pm 7)$ são isomorfos entre si e gerados por ação conjunta, e admitem um subgrupo de translações à direita com gerador γ^3 , que pode produzir a geodésica fechada mais curta para algum observador.

A presença ou não de holonomias translacionais depende tanto do valor de p quanto de q . Mas é imediato (c.f. D.10) que para um caso específico, quando p é primo, não existem translações admissíveis para qualquer valor de p .

5.2.2 Obtendo as geodésicas fechadas mais curtas

Uma condição necessária para a detectabilidade em princípio da topologia cósmica é que a geodésica fechada de menor comprimento na vizinhança do observador seja menor que o diâmetro do Universo observável. Assim, procuramos caracterizar qual é este comprimento para espaços-lente $L(p, q)$. Em cada ponto, a geodésica mais curta liga pares de pontos identificados por uma isometria $\gamma_{(p,q)}^n$ pertencente ao grupo cíclico gerado $\gamma_{(p,q)}$ para algum $n < p$. Lembrando das definições no Capítulo 3, chamamos de $r_{inj}(z)$ o raio da maior esfera com centro em z que pode ser inscrita no domínio fundamental, que é exatamente metade do comprimento da menor geodésica fechada que corta o ponto z . O raio de injetividade global, r_{inj} , é o valor mínimo de $r_{inj}(z)$ para toda a variedade. Para um espaço-lente (ou para um subgrupo cíclico das holonomias do espaço), podemos dizer a partir de (5.2) que $r_{inj}(z)$ é uma função somente de $|z_1|$. Para $|z_1| = 1$ ($|z_2| = 0$), temos a partir de (5.1) que $r_{inj}(z_1, 0) = \pi/p = r_{inj}$, que é claramente o menor raio de injetividade possível. De forma análoga, para $|z_1| = 1$ ($|z_2| = 0$), (5.3) garante que igualmente $r_{inj}(0, z_2) = \pi/p = r_{inj}$. Assim, para os valores extremos de $|z_1|$, $|z_1| = 0$ e $|z_1| = 1$, sabemos quais holonomias geram as geodésicas mais curtas: $\gamma_{(p,q)}$ e $\gamma_{(p,q)}^{\phi(p)-1}$. Veremos que as holonomias 'curtas' correspondentes aos valores intermediários são expressas em termos de inteiros que surgem naturalmente como denominadores de aproximações sucessivas de p/q , de uma forma que tornaremos explícita logo a seguir.

As ferramentas matemáticas necessárias e as provas são apresentadas no Apêndice B, mas para fixar a notação, podemos dizer sucintamente que qualquer número racional $\frac{q}{p}$, com $(q, p) = 1$, pode ser escrito como uma fração continuada finita

$$\frac{q}{p} = a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \ddots \frac{1}{a_{k-1} + \frac{1}{a_k}}}} , \quad (5.4)$$

onde $a_1 \in \mathbb{Z}$, $a_i \in \mathbb{N} \mid i > 1$ são os chamados quocientes parciais, a parte inteira de sucessivas aproximações da fração. Usaremos a representação $\frac{q}{p} \equiv [a_1, a_2, a_3, \dots, a_k]$, que podemos sempre definir de forma única. A fração $\frac{q}{p}$ pode ser sucessivamente aproximada pelos seus convergentes

$$c_i = \frac{q_i}{p_i} = [a_1, a_2, a_3, \dots, a_i] , \quad (5.5)$$

onde definimos ainda $c_k = q_k/p_k = q/p$.

Com estas definições em mente, as holonomias que geram as menores geodésicas fechadas são dadas por (c.f D.3)

Teorema 5.1 *Em um espaço-lente $L(p, q)$ a menor distância entre um ponto z e sua imagem por alguma holonomia $\gamma_{(p,q)}^n z \in \Gamma_{L(p,q)}$ é gerada pela isometria $\gamma_{(p,q)}^{p_j}$, onde p_j é o denominador irredutível de algum convergente de p/q .*

As geodésicas mais curtas nos equadores são casos especiais do teorema acima, e são geradas pelas holonomias $\gamma_{(p,q)}^{p_1}$ e $\gamma_{(p,q)}^{p_{k-1}}$, onde $p_1 = 1$ e (podemos provar, c.f. D.4), $p_{k-1} \equiv (-1)^k q^{\phi(p_k)-1} \pmod{p}$. Este último resultado fornece ainda uma maneira simples de obter a sequência dos p_i recursivamente a partir de p .

As geodésicas mais curtas podem ser sempre obtidas como indicado acima para qualquer p e q . Mas podemos também fazer afirmativas mais gerais sobre o comprimento de tais geodésicas em termos somente da ordem do grupo (c.f D.4), independente da posição do observador ou de q . Afirmamos

Teorema 5.2 *O raio de injetividade $r_{inj}^{L(p,q)}$ em um espaço-lente $L(p, q)$, é tal que $r_{inj}^{L(p,q)} \leq \frac{\pi}{\sqrt{p}}$.*

Definindo $r_{\max}$ como o raio de injetividade máximo na variedade, $r_{\max} = \max r_{inj}(z)$, escrevemos

$$r_{\max} \leq \frac{\pi}{\sqrt{p}} = r_{inj} \sqrt{p}. \quad (5.6)$$

Para valores suficientemente grandes de p , isto garante que em qualquer ponto da variedade existe pelo menos uma geodésica fechada 'curta' (i.e., pequena em relação ao raio de curvatura). Em formas espaciais esféricas com vários geradores, cada subgrupo cíclico irá gerar uma cota superior para $r_{\max}$ do tipo (5.6) em função da ordem do subgrupo. Assim, em uma forma espacial esférica genérica, tanto r_{inj} (que provê condições necessárias para a detectabilidade, c.f. o Capítulo 3) quanto $r_{\max}$ (que provê condições suficientes) são dados pela ordem do maior subgrupo cíclico das holonomias. Se este subgrupo é translacional, então obviamente $r_{\max} = r_{inj}$.

Ainda no tópico detectabilidade, assim como fizemos na Seção 3.3.4, se admitirmos um vínculo entre parâmetros de um modelo Λ CDM, (e.g. o dado em [158], pp. 58), podemos usar (5.6) para obter condições suficientes para a detectabilidade da topologia cósmica no caso de espaços-lente $L(p, q)$ (ou, de forma mais geral, de formas esféricas que incluam a isometria $\gamma_{(p,q)}$), em termos do parâmetro de densidade total Ω_0 . Os resultados

para alguns espaços-lente são mostrados na Tabela 5.2. Compare com os resultados da Tabela 3.4, onde são tabuladas as condições necessárias para a detectabilidade destas mesmas variedades. Note que a detectabilidade de nenhum dos espaços-lente selecionados é garantida para valores de Ω_0 admitidos em 2σ para as melhores observações disponíveis [158].

Espaço-lente	$r_{\max}$	Ω_0 min
$L(18, q)$	$\frac{\pi}{\sqrt{18}}$	1.124
$L(12, q)$	$\frac{\pi}{\sqrt{12}}$	1.205
$L(10, q)$	$\frac{\pi}{\sqrt{10}}$	1.263
$L(8, q)$	$\frac{\pi}{\sqrt{8}}$	1.363
$L(6, q)$	$\frac{\pi}{\sqrt{6}}$	1.587
$L(4, q)$	$\frac{\pi}{\sqrt{4}}$	2.50

Tabela 5.2: Valores mínimos para Ω_0 que garantem detectabilidade em qualquer ponto de uma seleção de espaços-lente, tomando o vínculo $\Omega_0 = 1,306 - 0,876 \Omega_{m0}$, que corresponde à direção de máxima degenerescência da região de confiança obtida a partir de dados do WMAP, no modelo Λ CDM.

Uma outra cota superior para $r_{\max}$, independente da posição mas dependente de q , pode ser obtida se restringirmos a análise a geodésicas fechadas obtidas a partir de $\gamma_{(p,q)}$. Uma consequência imediata de (5.2) é que o comprimento mínimo de tais geodésicas é sempre $\leq \frac{2\pi q}{p}$, e que portanto $r_{\max} \leq \frac{\pi q}{p}$. Esta cota é melhor que (5.6) se e somente se $q < \sqrt{p}$. Incidentalmente, neste caso $p_2 = \text{int}[\frac{p}{q}]$, e portanto $p_2 > \sqrt{p}$. Assim, pelo Teorema D.4, a holonomia que gera a mais curta geodésica fechada é $\gamma_{(p,q)}$ pelo menos na metade do domínio fundamental que corresponde ao 3-hemisfério $|z_2| \leq |z_1|$.

As cotas obtidas acima podem ainda ser usadas para separar o espaço paramétrico $p-q$ dos espaços-lente em regiões nas quais a topologia é detectável, indetectável, e excluída por observações. Mostramos a figura esquemática na Figura 5.1.

Para terminar de caracterizar a forma local dos espaços-lente, as holonomias associadas às geodésicas subsequentes em ordem de tamanho podem ser obtidas de maneira análoga à usada para obter a geodésica mais curta.

Lema 5.2 *Se em um espaço-lente $L(p, q)$ a menor distância entre um ponto z e sua imagem por alguma holonomia $\gamma_{(p,q)}^n z \in \Gamma_{L(p,q)}$ é dada pela isometria $\gamma_{(p,q)}^{p_j}$, então a segunda*

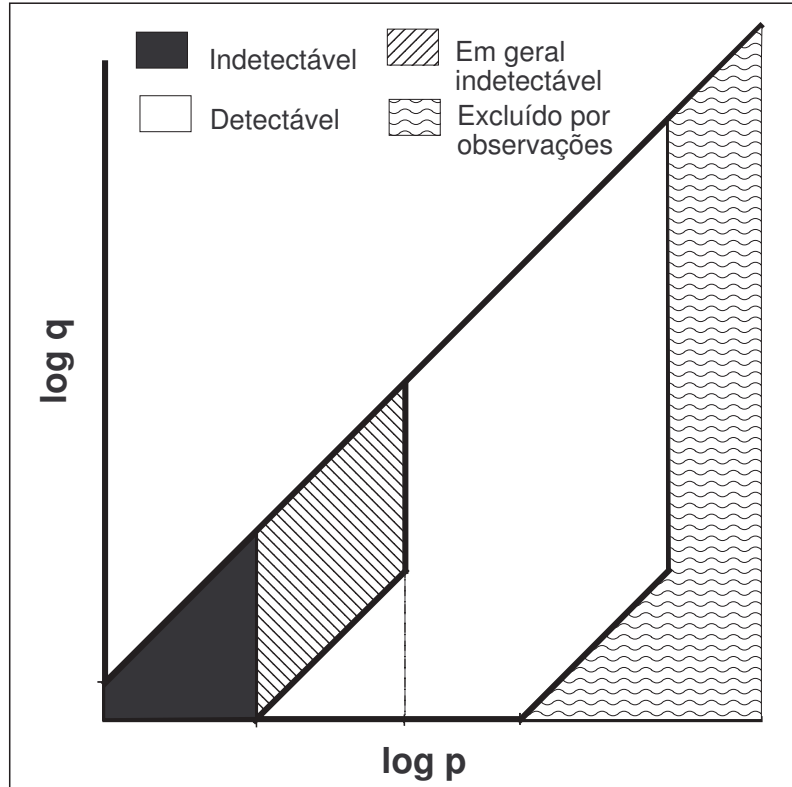


Figura 5.1: Detectabilidade de espaços-lente $L(p, q)$ em termos dos parâmetros p e q . Uma busca por repetição de padrões de origem topológica deve se concentrar na região branca. Variedades na região preta são excluídas por observações; na região ondulada a topologia é indetectável, e na região cinzenta a sua condição de detectabilidade varia caso a caso.

menor distância entre z e alguma de suas imagens topológicas é gerada por uma das seguintes isometrias: $\gamma_{(p,q)}^{p_i}$ para algum $i \neq j$, $\gamma_{(p,q)}^{2p_j}$ ou $\gamma_{(p,q)}^{p_{j+1}-p_j}$. Em particular, se $a_{j+1} = 1$ então se verifica a primeira alternativa.

O lema acima pode ser estendido trivialmente para as geodésicas fechadas subsequentes em ordem de comprimento. Assim, se para um ponto qualquer k é o número de possíveis geradores da menor geodésica, então serão no máximo $k + c - 1$ os possíveis geradores da c -ésima menor geodésica. Mostramos ainda que $k \leq 2 \log_2 2q$ (c.f. D.5), ou seja, que para um grupo cíclico de ordem alta, o número de possíveis candidatos a geradores das menores geodésicas é relativamente pequeno.

Finalmente, é interessante saber quando as holonomias que são translações de Clifford, que caracterizamos na subseção anterior, podem ser geradoras das geodésicas fechadas mais curtas em algum ponto. Provamos acima que uma condição necessária para seja este o caso para o gerador de um subgrupo translacional é que este seja a p_j -ésima iteração do gerador do grupo completo $\gamma_{(p,q)}$. O lema a seguir (c.f. D.11) explicita a condição necessária para que isto se verifique.

Lema 5.3 *Seja $\gamma_{(p,q)}$ um gerador de holonomias esféricas. Se $\gamma_{(p,q)}^n$ é uma translação de Clifford tal que $n \leq \sqrt{p/2}$, então $n = p_j$ para algum j .*

Na Tabela 5.1, translações de Clifford presentes nos grupos de holonomias dos espaços-lente que podem gerar a geodésica fechada mais curta são sublinhadas.

5.3 A forma local do Universo no limite inflacionário

5.3.1 A inflação

Repetindo um processo que parece ocorrer de forma recorrente na história da cosmologia, no final da década de 70 modelos do tipo Big Bang quente haviam tido grande sucesso em explicar a maior parte das observações disponíveis, e acreditava-se que o entendimento do cosmos na época era fundamentalmente correto, e razoavelmente completo. Havia porém algumas discrepâncias, coincidências que pareciam exigir condições iniciais precisamente definidas sem uma boa razão aparente (c.f. [94] para maiores detalhes). Em particular,

- Monopolos magnéticos, e outras relíquias da época dominada por radiação, não foram detectadas. Modelos de física de partículas tipicamente prevêem a sua produção em profusas quantidades em épocas primordiais.
- A radiação cósmica é quase isotrópica. Regiões situadas em direções diametralmente opostas do céu não deveriam ter tido tempo de entrar em contato causal até a época da recombinação, mas pareciam estar em perfeito equilíbrio térmico.
- O Universo é quase plano. Mesmo sem a precisão das observações atuais, já se sabia que a densidade total era da ordem de grandeza do valor crítico. O problema é que em um Universo em desaceleração, a quantidade $|\Omega(t) - 1|$ aumenta com o tempo. Isso significa que se $|\Omega_0 - 1| \sim \mathcal{O}(1)$ hoje, então $|\Omega(t) - 1| \lesssim \mathcal{O}(10^{-16})$ na época da nucleossíntese. É um número extremamente pequeno, e de fato valores muito maiores implicariam em um recolapso antes que a nucleossíntese pudesse ocorrer, para curvaturas positivas; ou uma expansão tão rápida que ela não seria capaz de seguir o seu curso, para curvaturas negativas. Não havia mecanismo plausível que determinasse com tamanha precisão o parâmetro de densidade inicial. Poder-se-ia obviamente supor que o Universo é exatamente plano, e $\Omega_0 = 1$ para todo o sempre, mas isto só tornava a coincidência mais excepcional.

De forma geral, todos estes problemas têm a mesma origem: o Universo observável oriundo da teoria clássica do Big Bang quente era grande demais. Grande demais para não conter monopolos, grande demais para ter estado em equilíbrio térmico, e grande demais para que efeitos da curvatura só fossem relevantes além do horizonte.

A inflação é, fundamentalmente, um mecanismo natural para torná-lo menor. Isto é visto mais claramente considerando a equação de Friedmann (1.12)

$$|\Omega(t) - 1| = \frac{|k|}{H^2 a(t)^2}; \quad (5.7)$$

O termo à direita combina as duas escalas características cosmológicas: A escala, ou tempo, de Hubble, H^{-1} (estamos tomando aqui novamente $c = 1$), e o raio de curvatura, a . A combinação dos dois produz H^{-1}/a , que podemos chamar de escala de Hubble comóvel, medida em unidades do raio de curvatura. Esta escala é da ordem de grandeza do raio do horizonte cosmológico, $\chi(t) = \int_{z(t)}^{\infty} H(z)^{-1}/a(t) dz$ (c.f. a seção 3.2). Mas também, crucialmente, $H^{-1}/a = \dot{a}^{-1}$. Assim, em um Universo dominado por matéria e radiação, a ação da gravidade faz com que a expansão desacelere, $\ddot{a} < 0$, e o raio comóvel do horizonte cosmológico cresce monotonicamente. Porém, se por alguma razão a expansão é acelerada, e $\ddot{a} > 0$, então a escala de Hubble comóvel decresce. À medida em que isto ocorre, pontos que até então faziam parte do Universo observável afastam-se cada vez mais rápido, até cruzarem o horizonte, e concomitantemente a densidade total vai se aproximando da densidade crítica. De um ponto de vista comóvel, o raio do horizonte cosmológico vai diminuindo. Se este processo inflacionário continua por tempo suficiente, a densidade de relíquias anteriores à inflação torna-se irrisória, e uma pequena região causalmente conectada se expande até além do horizonte. No final do processo, o raio do horizonte cosmológico fica tão pequeno em relação ao raio de curvatura que para todos os efeitos práticos a geometria (mas não, como sabemos, a topologia) do Universo observável torna-se efetivamente plana. A hipótese inflacionária consiste em postular uma tal fase de expansão acelerada.

A inflação foi proposta pela primeira vez por Alan Guth [69] e (em um contexto diferente) Alexei Starobinski [161] no início da década de 80, e desenvolvida em sua forma moderna inicialmente por Andrei Linde e colaboradores [96]. Existe atualmente um enorme corpo de literatura a respeito de modelos inflacionários e suas conseqüências observacionais (veja [94] para uma excelente introdução, e também [128]). Embora aspectos importantes da teoria, tais como a natureza do mecanismo que inicia e termina o processo inflacionário, tenham evoluído bastante desde sua concepção original, a intuição fundamental de uma fase acelerada de expansão dando origem a um Universo observável

(quase) plano e homogêneo permaneceu inalterada.

Modelos inflacionários postulam a existência na época anterior à nucleossíntese de (pelo menos) um campo escalar, homogêneo na pequena região que deu origem ao nosso Universo observável. Campos escalares surgem naturalmente na física de partículas na descrição de partículas de spin nulo (e.g. o arredio boson de Higgs), embora nenhum tenha sido ainda detectado, e não existe consenso sobre um candidato natural ao papel do 'inflaton'. Do ponto de vista da cosmologia, é possível definir uma pressão $p = \frac{1}{2}\dot{\phi}^2 - V(\phi)$ e densidade $\rho = \frac{1}{2}\dot{\phi}^2 + V(\phi)$ associadas à um campo escalar, identificando os termos do tensor energia-momento deste com os termos de um fluido perfeito. Para condições iniciais apropriadas, o termo potencial domina, a pressão se torna negativa, e a fase de expansão acelerada tem início. Muito rapidamente, a densidade de todos os demais componentes materiais se torna negligível, e a expansão entra em uma fase exponencial³.

A dinâmica do campo, dada pelas equações de Friedmann, implica que o termo potencial vai diminuindo, até o campo terminar por oscilar em torno de um mínimo, decaindo em matéria convencional através de acoplamentos com outros campos, quando então o processo inflacionário tem fim. Sob as condições normalmente impostas sobre o potencial $V(\phi)$, ditas de 'slow roll',⁴ a expansão dura tempo suficiente para eliminar os problemas indicados acima, mas termina naturalmente de modo que o Universo possa ser dominado por radiação, com flutuações de densidade capazes de, muito tempo depois, formar estruturas habitadas por cosmólogos inquisitivos. Tipicamente, isto implica que a região do Universo sofrendo inflação se expande por um intimidador fator de pelo menos e^{70} . A escala de Hubble medida em unidades do raio de curvatura diminui pelo mesmo fator, o que implica que, para valores pré-inflacionários razoáveis, atualmente teríamos (c.f. [97] e [163])

$$|\Omega_0 - 1| \lesssim 10^{-5} \ll 1. \quad (5.8)$$

No que se segue, chamaremos a desigualdade acima de limite inflacionário.

Não dispomos atualmente de meios de testar a validade do limite inflacionário. A medição da posição do primeiro pico acústico do espectro de potência angular da radiação cósmica de fundo, especialmente quando combinada com outras observações, impôs severas restrições ao valor do parâmetro de densidade total, Ω_0 . Mesmo assim, com uma

³Modelos com uma expansão do tipo lei de potência já foram propostos, mas são considerados um tanto exóticos.

⁴O nome tem origem em uma das equações de Friedmann, que toma a mesma forma da equação de uma partícula amortecida em um poço de potencial gravitacional; estas condições garantem que o termo dissipativo domine por um período suficientemente longo para que a inflação ocorra satisfatoriamente.

confiança de 2σ , temos $\Omega_0 = 1,014 \pm 0,17$, que é uma precisão pelo menos três ordens de grandeza aquém do necessário. Só podemos afirmar, com base em observações diretas, que o Universo é plano ou quase plano; como vimos no Capítulo 3, como estas observações ainda admitem curvaturas das seções espaciais do Universo positivas e negativas, não há limite superior para o valor do raio de curvatura, o que implica que qualquer topologia não-plana (cujas geodésicas fechadas tem sempre comprimentos invariantes quando medidos em unidades do raio de curvatura) pode ser potencialmente indetectável para uma escolha admissível dos parâmetros de densidade.

Além de determinar com precisão inédita Ω_0 , porém, o WMAP também confirmou indiretamente muitas das previsões de modelos inflacionários, tais como perturbações de densidade adiabáticas, gaussianas e quase invariantes por escala (veja a Seção 1.3.3 e o Apêndice A). Existem algumas discrepâncias (como e.g. indícios de não gaussianidade, [44]) ainda a serem explicadas. Mas por enquanto o paradigma inflacionário é aceito como parte do modelo padrão da cosmologia. Faz sentido portanto tomá-lo como hipótese, para estudar a forma local do Universo. É o que faremos a seguir. Vamos supor que o limite inflacionário valha, e que o Universo seja não plano, e dotado de uma topologia não trivial detectável.

5.3.2 A forma local do Universo no limite inflacionário

Podemos expressar o limite inflacionário em termos das quantidades que determinam a detectabilidade da topologia (c.f. Capítulo 3), o raio de injetividade local $r_{inj}(x)$ (onde x é a posição do observador), e a profundidade de um levantamento, χ_{obs} . Qualitativamente, supondo que a topologia é detectável, temos que, em unidades do raio de curvatura,

$$r_{inj}(x) \lesssim \chi_{obs} \ll 1. \quad (5.9)$$

Como vimos, podemos usar a reta secante para obter condições suficientes para detectabilidade, o que equivale a dizer que podemos usá-la para obter uma cota inferior de χ_{obs} . A partir de (3.14) obtemos assim⁵

$$\chi_{obs} \lesssim 2\sqrt{\frac{|\Omega_0 - 1|}{\Omega_{m0}}}, \quad (5.10)$$

onde usamos a hipótese de que $r_{inj}(x) \ll 1$ para escrever $\sin^2(r_{inj}^M/2) \simeq r_{inj}^M/4$

⁵No limite inflacionário, a diferença entre as curvas de contorno $\chi_{obs} = r_{inj}^M(x)$ para variedades M detectáveis é, para todos os efeitos práticos, indistinguível da reta secante (que, como vimos, já é uma boa aproximação mesmo para $|\Omega - 1| \simeq 0,02$); assim, podemos trocar a desigualdade $<$ por $\lesssim$.

Combinando o resultado acima com (5.8) e (5.9), obtemos uma expressão mais direta do limite inflacionário,

$$r_{inj}(x) \lesssim \chi_{obs} \lesssim \frac{2}{\sqrt{\Omega_{m0}}} 10^{-5/2} \simeq 0.01 , \quad (5.11)$$

onde o valor final foi obtido assumindo (assim como fizemos no Capítulo 3) que $\Omega_{m0} = 0,35$.

Vamos a seguir considerar holonomias detectáveis no limite inflacionário, tanto para curvaturas positivas quanto negativas, e em particular determinar o quão próximo sua ação restrita ao Universo observável é a de uma translação de Clifford. Para tanto, vamos obter cotas para a variação da distância entre pares de pontos identificados pela isometria que se situam no interior de uma esfera de raio χ_{obs} .

Variedades esféricas

Mostramos anteriormente (5.6) que variedades esféricas em geral possuem holonomias 'curtas'. De forma mais precisa, os subgrupos cíclicos de ordem p do grupo de holonomia geram em qualquer ponto geodésicas fechadas de comprimento menor que $\pi/\sqrt{p}$. Isto significa que para variedades com subgrupos cíclicos de ordem suficientemente alta, a topologia é detectável em qualquer ponto.

Se uma holonomia β é detectável em uma vizinhança de um ponto z , então a geodésica fechada associada é tal que $d(z, \beta z) \ll 1$. Supondo que seja esta a geodésica fechada mais curta, então sabemos que $\beta = \gamma^{p_j}$, para algum j , e onde γ é o gerador do subgrupo cíclico. Podemos então expandir (5.2), e obter

$$d(z, \gamma^{p_j} z) \simeq 2\pi \left[\left(\frac{p_j}{p} \right)^2 |z_1|^2 + \left(\frac{1}{p_{j+1} + p_j/\alpha_{j+2}} \right)^2 |z_2|^2 \right]^{1/2} , \quad (5.12)$$

acima, $\alpha_i = [a_i, a_{i+1} \dots a_k]$.

Usando o fato de que $|z_1|^2 + |z_2|^2 = 1$ obtemos

$$d(z, \gamma_{(p,q)}^{p_j} z) \simeq 2\pi [H_j^2 + (G_j^2 - H_j^2) |z_2|^2]^{1/2} , \quad (5.13)$$

onde $H_j = p_j/p$ e $G_j = (p_{j+1} + p_j/\alpha_{j+2})^{-1}$. A fórmula acima pode ser entendida como a função distância da isometria $\gamma_{(p,q)}^{p_j}$ no ponto z . Deixando implícita a dependência da isometria, usaremos a notação $d_z = d(z, \gamma^{p_j} z)$. Consideramos agora dois casos:

Caso I. $|z_1|$ e $|z_2| \gg \chi_{obs}$.

Pela fórmula (5.13), vemos que a variação máxima da função distância entre um ponto z e sua imagem por p_j , se dá quando a variação de $|z_2|$ é máxima. É preciso, portanto,

determinar a variação máxima de $|z_2|$ em uma esfera de raio χ_{obs} centrada no ponto z . Seja então $w \in \mathbb{S}^3$ um outro ponto tal que $d(w, z) = \chi_{obs}$. Seja $\cos \zeta = |z_2|$ (de modo que $\sin \zeta = |z_1|$). Claramente, $\cos(\zeta + \chi_{obs}) \leq |z_2 + w_2| \leq \cos(\zeta - \chi_{obs})$. Lembrando que estamos assumindo (5.9), temos $|z_2| - |z_1|\chi_{obs} \leq |z_2 + w_2| \leq |z_2| + |z_1|\chi_{obs}$, e como consequência

$$d_{\pm} \simeq 2\pi \left[H_j^2 + |G_j^2 - H_j^2| |z_2|^2 \left(1 \pm \frac{|z_1|}{|z_2|} \chi \right)^2 \right]^{\frac{1}{2}}, \quad (5.14a)$$

$$d_{\pm} \simeq d_z \pm (2\pi)^2 |G_j^2 - H_j^2| \frac{|z_1||z_2|}{d_z} \chi. \quad (5.14b)$$

Note que $|G_j^2 - H_j^2| |z_1|^2 |z_2|^3 > G_j^2 |z_1|^2 + H_j^2 |z_2|^2 = d_z$. Assim, resulta que

$$d_{\pm} = d_0 \pm (2\pi)^2 |G_j^2 - H_j^2| \frac{|z_2||z_1|}{d_0} \chi_{obs} \quad (5.15)$$

A diferença entre os comprimentos máximo e mínimo das geodésicas fechadas contidas na esfera de raio χ_{obs} , e geradas pela isometria γ^{pj} é então

$$\frac{\Delta d}{d_0} \leq 2 \frac{1}{|z_2||z_1|} \chi_{obs} \ll 1. \quad (5.16)$$

Esta é uma cota para a variação máxima do comprimento de geodésicas fechadas geradas pela mesma holonomia no Universo observável (Veja a Figura 5.2). Claramente, com estas hipóteses, a isometria é muito semelhante a uma translação de Clifford. Em particular, um par de círculos no céu ligado por esta isometria será quase antipodal. Se tais círculos tem centros com coordenadas $C_1 = (\theta_1, 0)$ e $C_2 = (\theta_2, 0)$, então é possível mostrar que o desvio em relação à antipodicidade, $\pi - C_1.C_2 \simeq \frac{\Delta d}{2\chi_{obs}} \leq \frac{\Delta d}{d_0}$. Assim, a busca por pares de círculos no céu, computacionalmente bastante trabalhosa (veja a discussão na Seção 4.2), é bastante simplificada se tomarmos como hipótese o limite inflacionário, pois o espaço paramétrico a ser investigado se reduz sensivelmente. Uma restrição semelhante foi usada nas buscas por círculos no céu empreendidas por Cornish *et al.* [39] e Shapiro Key *et al.* [151], que não encontraram pares de círculos correlacionados.

Caso II. $|z_1|$ ou $|z_2| \sim \chi_{obs}$.

Se $|z_1| \sim 1$ (o caso $|z_2| \sim 1$ é idêntico), então podemos assumir sem perda de generalidade que a menor geodésica é gerada por $\gamma_{(p,q)}$. Obtemos assim

$$\begin{aligned} d_z^2 &\simeq 4r_{inj}^2 + 2[1 - \cos(\frac{2\pi q}{p})]z_2^2, \\ d_{\pm}^2 &\simeq 4r_{inj}^2 + 2[1 - \cos(\frac{2\pi q}{p})](z_2 \pm \chi_{obs})^2. \end{aligned} \quad (5.17)$$

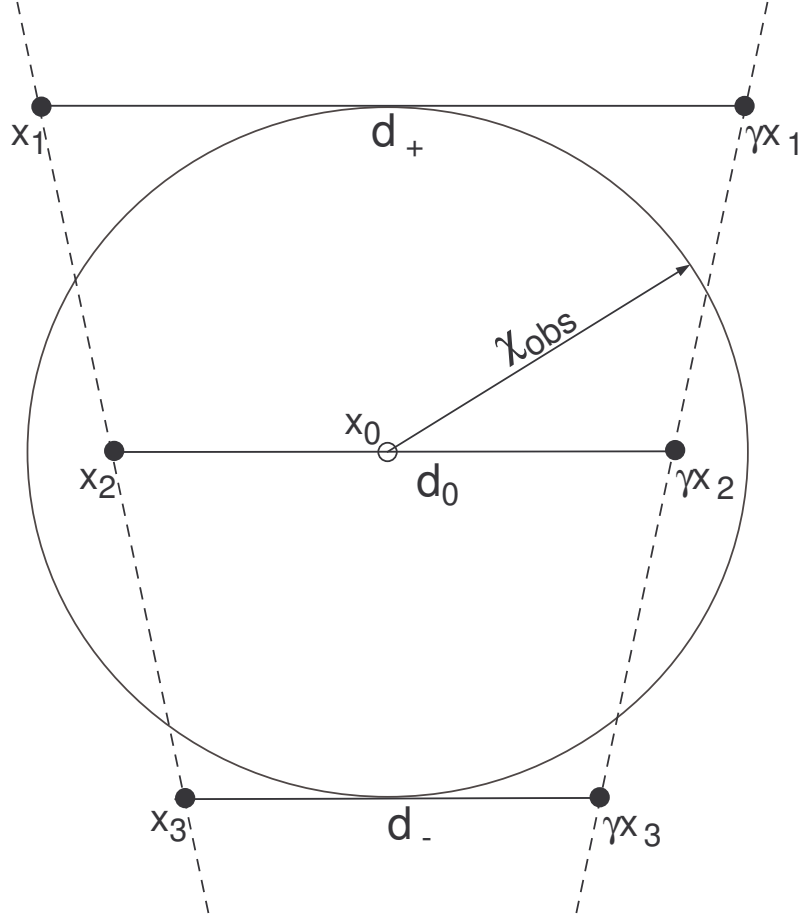


Figura 5.2: Representação esquemática das geodésicas fechadas mais curtas (comprimento d_-) e mais longa (d_+) geradas pela holonomia γ que interceptam o Universo observável (uma esfera de raio χ_{obs}). Estes valores são cotas respectivamente superior e inferior para as geodésicas fechadas detectáveis de comprimento máximo e mínimo geradas por γ .

Note que em ambas as equações os dois termos são da mesma ordem. Fica claro, portanto, que neste caso a isometria não se parece localmente com uma translação de Clifford. Mas no limite inflacionário o volume do espaço em que $|z_i| \sim \chi_{obs}$ é bastante reduzido. Se V_R é o volume desta região (para $i = 1$ e $i = 2$), e $V_{\mathbb{S}^3}$ o volume da 3-esfera, é fácil obter, usando (5.11)

$$\frac{V_R}{V_M} \sim \frac{3}{2} \chi_{obs}^2 \lesssim \mathcal{O}(10^{-4}), \quad (5.18)$$

que claramente é muito pequeno. Desta maneira, no limite inflacionário a probabilidade de um observador estar situado de maneira a detectar uma isometria que não se assemelhe a uma translação de Clifford é pequena, no sentido usado por Weeks em [182].

Variedades hiperbólicas

Embora não haja uma classificação completa das 3-variedades compactas hiperbólicas, existem cotas inferiores para o seu volume [2]. Tipicamente, as suas geodésicas fechadas têm comprimento da ordem de $V^{1/3}$, de modo que para valores suficientemente pequenos de χ_{obs} , a maioria dos observadores na maior parte das variedades não será capaz de detectar quaisquer holonomias; a forma local do Universo será em geral trivial. Estas afirmativas são precisadas em termos numéricos em [182].

Existe assim uma classe de variedades hiperbólicas não compactas, as chamadas variedades cuspidais, nas quais existem subregiões onde $r_{inj}(x)$ é arbitrariamente pequeno, para uma escolha apropriada de x . Estas regiões, de formato cônico e denominadas cúspides, assintoticamente podem ser dotados da métrica

$$ds^2 = \frac{dx_1^2 + dx_2^2 + dx_3^2}{x_3^2} ; x_3 > 1 . \quad (5.19)$$

Variedades cuspidal são ainda dotadas de um subgrupo abeliano de holonomias com geradores α e β , que na cúspide tomam a forma onde

$$\begin{aligned} \alpha x_i &= x_i + ax_1 ; a \in \mathbb{R} , \\ \beta x_i &= x_i + bx_2 ; b \in \mathbb{R} . \end{aligned} \quad (5.20)$$

Note que uma cúspide se parece com um 2-toro que vai se afunilando na direção x_3 , e onde o comprimento da menor geodésica fechada que passa por x , $2r_{inj}(x)$, é a distância entre x e αx (onde supomos sem perda de generalidade que $a \leq b$), e é obtida por integração de (5.19),

$$r_{inj}(x) = \frac{a}{x_3} . \quad (5.21)$$

Veja que a forma local destas variedades é determinada somente pelos valores de a e b .

As variedades cuspidais são ainda o limite de seqüências de variedades compactas de volume crescente. Isto significa que existem sempre infinitas variedades compactas com uma região com holonomias e métrica dadas aproximadamente por (5.19) e (5.20) com qualquer precisão desejada. É somente nestas variedades (que chamamos quase-cuspidal), e nelas somente nas regiões que têm cusps como limite, que é necessário investigar as holonomias detectáveis no limite inflacionário.

Considere agora a função distância entre um ponto e sua imagem por α , $d_{x_0} = d(x_0, \alpha x_0)$. Imagine agora um ponto x' situado na esfera de raio χ_{obs} centrada em x_0 : $d(x_0, x') = \chi_{obs}$. Procuramos obter os valores de x' que respectivamente maximizam e minimizam a função distância $d_{x'}$. Definimos $d_- = \min d(x'_0, \alpha x')|_{d(x_0, x')=\chi_{obs}} = d(x_-, \alpha x_-)$

e $d_+ = \max d(x'_0, \alpha x')|_{d(x_0, x')=\chi_{obs}} = d(x_+, \alpha x_+)$. Claramente, por (5.19), a variação da função d_x é máxima na direção $\hat{x}_3$, e $x_\pm = \pm|x_\pm|\hat{x}_3$. Neste caso, podemos integrar a métrica ao longo do eixo x_3 para obter

$$\begin{aligned}\chi_{obs} &= d(x_0, x_\pm) = \pm \int_{x_{30}}^{x_\pm} \frac{dz}{z} , \\ \chi_{obs} &= \log\left[\frac{x_\pm}{x_{30}}\right] , \\ x_\pm &= e^{\pm\chi_{obs}} x_{30} \simeq x_{30}(1 \pm \chi_{obs}) ,\end{aligned}\tag{5.22a}$$

onde na última linha utilizamos a hipótese do limite cosmológico.

A partir de (5.21) temos ainda

$$d_\pm = \frac{a}{x_\pm} d_{\max} ,\tag{5.23a}$$

e, portanto,

$$d_\pm \simeq \frac{a}{x_{30}}(1 \mp \chi) = d_{x_0}(1 \mp \chi) ;\tag{5.24}$$

finalmente, obtemos

$$\frac{d_+ - d_-}{d_{x_0}} \equiv \frac{\Delta d}{d_{x_0}} = 2\chi .\tag{5.25}$$

Note a semelhança com a fórmula análoga obtida para espaços esféricos (5.16). Novamente, toda holonomia detectável no limite inflacionário é localmente semelhante a uma translação de Clifford. Combinando os dois resultados, podemos afirmar que esta é uma característica geral da forma local do Universo para variedades não-planas no limite cosmológico. É um resultado geral, obtido a partir de uma única hipótese simples e bastante razoável. Podemos dizer ainda mais. É claro que no caso de variedades hiperbólicas, as holonomias quase translacionais terão um ou no máximo dois geradores independentes (α e β em (5.20)). No caso de espaços-lente, considere que a geodésica fechada mais curta, gerada por uma holonomia γ , tem $r_{inj}(x) \leq 2\pi/\sqrt{p}$ (5.6). Se, além de γ , existirem mais duas holonomias quase translacionais independentes, γ' e γ'' , o domínio fundamental será um quase toro, as geodésicas fechadas por elas geradas deverão ser todas da mesma ordem de grandeza, o que implica que o volume deste domínio fundamental deveria ser $V_{L(p,q)} \sim V_{\mathbb{S}^3} p^{-3/2}$. Mas sabemos que o volume do domínio fundamental é na realidade $V_{L(p,q)} = V_{\mathbb{S}^3} p^{-1}$. Assim, por contradição, no limite inflacionário espaços-lente têm no máximo dois geradores independentes de quase translações. Mais geralmente, podemos afirmar que variedades não-planas no limite inflacionário são indistinguíveis (para $|\Omega_0 - 1|$ pequeno o suficiente) das variedades planas $\mathcal{E}$, $\mathcal{J}_0$ ou $\mathcal{T}_1$ (c.f. Apêndice B).

5.3.3 Consequências observacionais

Para métodos de repetição de padrão, as holonomias detectáveis são aquelas que geram localmente geodésicas fechadas de comprimento menor que o diâmetro do Universo observável, $2\chi_{obs}$. Assim, em variedades que não são globalmente homogêneas, o conjunto de holonomias detectáveis, que denominamos forma local do Universo, depende de forma crucial tanto da posição do observador quanto de χ_{obs} . Mas, como vimos, no limite inflacionário $\chi_{obs} \ll 1$, a forma local se simplifica enormemente. Isto tem consequências importantes sobre estratégias de busca da repetição de padrões, permitindo tornar tais buscas muito mais eficientes restringindo o espaço paramétrico a ser investigado.

O limite inflacionário também tem consequências significativas para testes indiretos da topologia cósmica, como por exemplo estudos de anomalias no espectro angular de potência da radiação cósmica de fundo. Em primeiro lugar, variedades com r_{inj} relativamente grande, como o dodecaedro de Poincaré, não são detectáveis diretamente no limite inflacionário, e teriam um efeito desprezível [142] em possíveis medidas indiretas. Além disso, o número de potenciais topologias a serem testadas cresce significativamente neste limite (com grupos cíclicos de ordens mais altas, no caso esférico, e com variedades que se aproximam assintoticamente de variedades cuspidais, no caso hiperbólico). Veja a Figura 5.1 para uma ilustração. O valor que usamos para quantificar o limite inflacionário, (5.8), implica em que variedades esféricas detectáveis em princípio devem ter um subgrupo cíclico de ordem $p \gtrsim 100$, e somente subgrupos com ordem $p \gtrsim 10000$ são detectáveis em todo o espaço. Lembramos que, para cada valor de p , existem aproximadamente $p/2$ subgrupos possíveis (ou espaços-lente), cada qual com p isometrias. É inconcebível testar todas as combinações de p , q e posição do observador possíveis simultaneamente. Torna-se então necessário lançar mão de um entendimento teórico que nos permita fazer afirmações mais gerais, ao invés de testar individualmente cada alternativa possível. É isto que procuramos fazer neste capítulo. Para variedades esféricas, mostramos como obter a forma local em termos das holonomias mais 'curtas' para um dado valor de p e q , cujo valor sabemos calcular. Em geral, subgrupos cíclicos diferentes podem ter formas locais semelhantes. Assim, uma maneira eficiente de se estudar os efeitos da topologia cósmica no espectro de potência angular seria determinar tal efeito em função da forma local, e não do grupo completo de holonomias. Se o espectro assim calculado não for compatível com o espectro observado, poderemos então excluir classes inteiras de topologias. Se por outro lado os espectros coincidirem, teremos uma forte indicação de que a topologia do Universo pertence a uma classe bem definida.

Este último ponto merece ser enfatizado. Em geral se assume que a topologia possa

ser determinada de forma unívoca. Mas, como vimos, no limite inflacionário, diferentes topologias têm uma forma local essencialmente idêntica. Assim, da mesma maneira que existe para todos os efeitos práticos uma degenerescência geométrica na medição do sinal da curvatura do espaço, poderemos nos ver em uma situação em que exista uma degenerescência topológica que não permita determinar, dentro de uma dada classe, qual exatamente é a topologia do Universo.

Uma classe importante de holonomias consiste nas translações de Clifford. Métodos de detecção da topologia cósmica que são sensíveis somente a este tipo de holonomia introduzem um grau adicional de degenerescência, pois variedades diversas muitas vezes têm os mesmos subgrupos translacionais.

Em um plano mais especulativo, seria interessante, por exemplo, investigar se espaços-lente que não admitem translações de Clifford como holonomias geram espectros de potência sistematicamente diferentes daqueles gerados por espaços-lente com subgrupos translacionais. De forma mais geral, a estrutura dos fatores primos, assim como a seqüência de convergentes p_i/q_i para a razão p/q , varia radicalmente para valores muito próximos de p (ou q). É razoável supor que isto se reflita no espectro de potência.

O estudo da forma local do Universo ainda é por enquanto incipiente. Mas já foi possível obter alguns resultados bastante gerais, o que dá razões para otimismo quanto às suas perspectivas, como uma nova abordagem para o problema da detecção e detectabilidade da topologia cósmica.

Capítulo 6

Efeito do movimento próprio de nossa galáxia na detecção de círculos no céu

Acreditar é tedioso. Duvidar é absorvente. Permanecer alerta é viver, deixar-se cair em um senso de falsa segurança é a morte.

Oscar Wilde

6.1 Introdução

Se levarmos em conta a qualidade e alta resolução dos mapas disponíveis da radiação cósmica de fundo (RCF), comparadas com as incertezas na determinação da posição de galáxias e outros objetos pontuais, fica claro que a maneira mais promissora de procurar uma possível topologia cósmica é pela detecção de repetições de padrões nestas anisotropias da RCF. Se uma topologia não trivial existe e é detectável (c.f. Capítulo 3), a última superfície de espalhamento (USE) ficará em todo ou em parte fora do domínio fundamental, e portanto interceptará algumas de suas imagens (veja a discussão no Capítulo 4). Para observadores comóveis, a USE é uma esfera (se desprezarmos os efeitos de sua espessura), e portanto estas intersecções serão circunferências. Assim, os mapas das anisotropias de temperatura na RCF deverão conter pares de círculos correlacionados de mesmo raio, com o mesmo padrão de variações de temperatura [39] na sua borda. Estes pares de círculos devem existir, distribuídos em diferentes padrões, para qualquer topolo-

gia não trivial detectável. Assim, para estudar a topologia do Universo na maior escala acessível através de observação de radiação eletromagnética, torna-se necessário usar uma abordagem estatística que análise mapas de céu inteiro da RCF a procura de pares correlacionados de círculos. O método de círculos no céu, proposto por Cornish, Spergel e Starkman em 1998 [39], introduz uma função de correlação de dois pontos, definida no espaço de parâmetros dos círculos, que é capaz de detectar tais pares consistentemente em simulações.

Da maneira como foi concebido originalmente, o método de círculos no céu não leva em conta o movimento próprio da nossa galáxia. Este ponto foi levantado pela primeira vez por Levin [82], que demonstra (no caso específico de um espaço-tempo plano) que para um observador não comóvel a USE, esférica para observadores comóveis, seria deformada de modo que as interseções da USE com suas imagens topológicas (doravante, os círculos no céu, ou simplesmente círculos quando não houver possibilidade de confusão) sofreriam deslocamento e deformação. É importante lembrar que as observações das flutuações da RCF obviamente não incluem informação direta sobre a distância das flutuações que se situam na USE, somente a sua distribuição angular; o que observamos é portanto a projeção da USE na esfera celeste. Assim, é esta projeção que é o resultado efetivamente relevante no referente a comparações com dados observacionais.

Como mostraremos abaixo, a interseção da USE com suas imagens topológicas, quando projetada na esfera celeste de um observador em movimento relativo ao referencial comóvel local, mantém-se circular para qualquer velocidade do observador. Este resultado é obtido a partir de medições puramente locais, e portanto é válido para qualquer valor da curvatura do 3-espaco, e independente da expansão do Universo.

De modo geral, observadores diferentes detectarão, associados a cada uma destas interseções, círculos de diferentes raios e centros, e com diferentes distribuições de anisotropias ao longo de seu perímetro. Na seção seguinte iremos obter expressões analíticas para todos estes três efeitos: a variação do raio, a variação da posição do centro, e o deslocamento angular, que juntos descrevem completamente as mudanças sofridas por círculos na esfera celeste devidas ao movimento próprio do observador.

6.2 A aberração local da radiação cósmica

Um teorema demonstrado por Penrose em 1959 [130] mostra que se um observador O detecta um objeto circular (quando projetado na esfera celeste), então qualquer outro observador O' em movimento relativo e posição coincidente a O também detectará a

forma do objeto como circular.¹

No que nos concerne, o importante corolário deste resultado é que círculos no céu, que são definidos em princípio somente para observadores comóveis, serão detectados como círculos por quaisquer observadores. Mas os pares de círculos identificados por isometrias, que são idênticos para o observador comóvel, terão em geral raios e o padrão circular de anisotropias diferentes, e sua posição relativa também poderá ser alterada. Nos resta então quantificar estes efeitos de aberração.

Seja $P^\mu = (h\nu/c)(1, -\hat{n})$ o 4-momento de um fóton vindo da direção $\hat{n}$ com frequência ν , detectado pelo observador O em repouso em relação ao fluxo de Hubble (um observador comóvel). Para um segundo observador O' , em movimento com velocidade $\vec{\beta} = \beta\hat{\beta}$ em relação ao primeiro, cuja posição espacial coincide com a de O quando ambos fazem observações da RCF, o mesmo fóton seria detectado com vindo de uma direção diferente, $\hat{n}'$, e com uma frequência diferente, ν' , dados por

$$\nu' = \gamma (1 + \vec{\beta} \cdot \hat{n}) \nu, \quad (6.1)$$

$$\hat{n}' = \frac{\hat{n} + [(\gamma - 1)\hat{\beta} \cdot \hat{n} + \gamma\beta] \hat{\beta}}{\gamma (1 + \vec{\beta} \cdot \hat{n})}. \quad (6.2)$$

A equação (6.1) nada mais é do que o efeito Doppler, ou seja, a mudança da frequência observada devido ao movimento relativo entre o observador (situado na Via Láctea) e o emissor (na USE), e é responsável pela quase totalidade do componente de dipolo da RCF, e por esta razão geralmente é removido dos mapas antes da análise das anisotropias. O efeito Doppler causa uma pequena distorção nos multipolos superiores, mas estes efeitos são desprezíveis (para uma análise detalhada dos efeitos do movimento próprio do observador na detecção de radiação térmica, veja [152]).

A equação (6.2), por outro lado, expressa a aberração relativística da luz, indicando que a radiação incidente será detectada como vindo de diferentes direções para os observadores O' e O . É este efeito que iremos estudar daqui em diante.

Primeiramente usamos (6.1) para determinar o centro e a posição de um círculo no céu genérico quando mudamos de um referencial para o outro. Sem perda de generalidade, escolhemos os eixos coincidentes z e z' de tal maneira que sua direção coincida com a direção da velocidade $\vec{\beta}$. Assim, a transformação que rege a mudança da direção de um

¹Este é um caso especial do chamado efeito Terrel-Penrose. Para um tratamento completo, veja por exemplo [104].

fóton incidente, $(\theta, \varphi) \mapsto (\theta', \varphi')$, é obtida diretamente de (6.1), e tem a forma

$$\cos \theta' = \frac{\beta + \cos \theta}{1 + \beta \cos \theta} \quad \text{e} \quad \varphi' = \varphi, \quad (6.3)$$

onde θ e φ são as coordenadas esféricas usuais.

Sejam $\hat{q}_c = (\theta_c, \varphi_c)$ e ρ respectivamente o centro e o raio do círculo $C(\hat{q}_c, \rho)$ no referencial comóvel. Note que a mudança de referencial dada por (6.3) não leva em geral o centro $\hat{q}_c$ de C no centro $\hat{q}'_c$ do círculo observado pelo observador em movimento, $C'(\hat{q}'_c, \rho')$. Para obter a posição correta de $\hat{q}'_c = (\theta'_c, \varphi'_c)$ e o raio ρ' de C' a maneira mais simples é aplicar (6.3) nos extremos superior e inferior do círculo C , dados respectivamente por $(\theta_c - \rho, \varphi_c)$ e $(\theta_c + \rho, \varphi_c)$. Da simetria do problema, como $\varphi' = \varphi$, então $\varphi'_c = \varphi_c$. Assim, os pontos obtidos também serão os extremos superior e inferior do círculo C' , e o centro $\hat{q}'_c$ é facilmente obtido como o ponto médio (ao longo da esfera celeste) entre eles. Da mesma maneira, o raio ρ' nada mais é que metade da distância angular entre estes dois pontos. Assim, aplicando relações trigonométricas básicas, podemos obter após algumas manipulações

$$\sin \theta'_c = \frac{1}{\gamma M} \sin \theta_c, \quad (6.4a)$$

$$\sin \rho' = \frac{1}{\gamma M} \sin \rho, \quad (6.4b)$$

onde

$$M = \sqrt{[1 + \beta \cos(\theta_c + \rho)][1 + \beta \cos(\theta_c - \rho)]}. \quad (6.5)$$

Na Figura 6.1 indicamos esquematicamente o resultado das Equações (6.4a) e (6.4b). São mostrados os raios e as posições dos centros dos círculos para três casos especiais. Traços cheios e tracejados representam os círculos no céu detectados pelos observadores O e O' respectivamente.

Com as equações (6.3) e (6.4), é fácil confirmar o resultado de Penrose. Com efeito, podemos definir o círculo $C(\hat{q}_c, \rho)$ como o conjunto de pontos $\hat{q}$ na esfera que obedecem a equação

$$\hat{q} \cdot \hat{q}_c = \cos \rho. \quad (6.6)$$

Verifica-se que a igualdade é preservada pela aplicação das transformações (6.3) e (6.4).

Expandindo a equação acima até a primeira ordem em β , o que é válido para velocidades não-relativísticas ($|\beta| \ll 1$), obtemos ainda

$$\theta'_c = \theta_c - \beta \sin \theta_c \cos \rho, \quad (6.7a)$$

$$\rho' = \rho - \beta \cos \theta_c \sin \rho. \quad (6.7b)$$

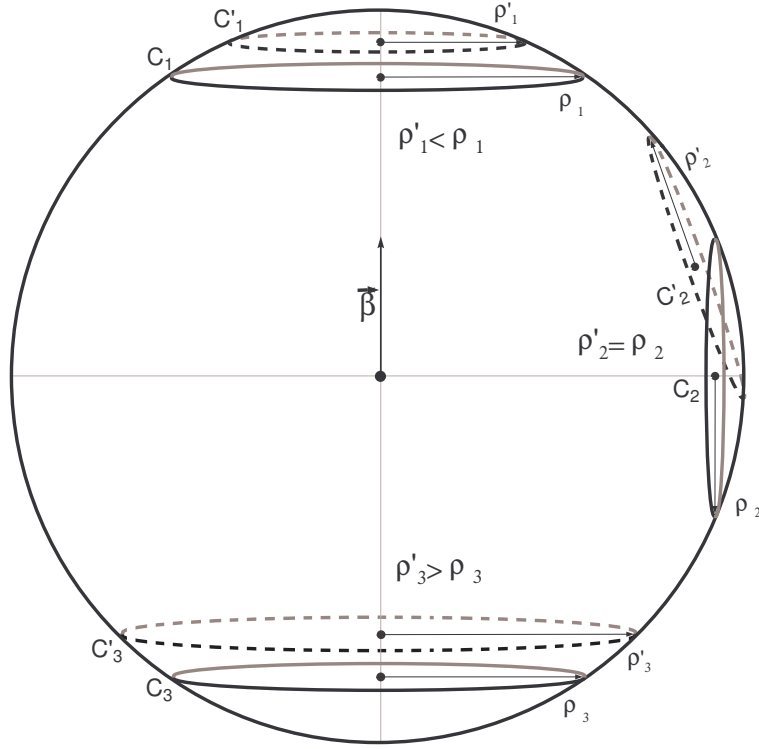


Figura 6.1: Ilustração esquemática da variação dos raios e posições dos centros dos círculos. Traços contínuos e tracejados indicam respectivamente os círculos no céu detectados pelos observadores O e O' . Os círculos C_2 e C'_2 mostram a variação máxima na posição do centro; nos casos em que isto acontece o raio permanece invariante ($\rho_2 = \rho'_2$). Os círculos C_1 e C'_1 e C_3 e C'_3 são os pares para os quais o raio do círculo sofre respectivamente diminuição e aumento máximos. Por simetria, nestes casos a posição do centro dos círculos não muda.

Como consequência de (6.7), é fácil ver que para um círculo de raio ρ o deslocamento de seu centro, $|\Delta\theta_c| = |\theta'_c - \theta_c|$, é máximo para $\theta_c = \pi/2$, i.e. quando a posição do centro do círculo no referencial comóvel é ortogonal à direção de movimento $\vec{\beta}$. Neste caso o raio do círculo é o mesmo para ambos os observadores. A cota superior de $|\Delta\theta_c|$ é $|\beta|$. Assim, dois círculos originalmente antipodais no referencial comóvel serão observados com os centros em uma posição deslocada em relação à posição antipodal por no máximo $2|\beta|$. A figura 6.2 mostra a variação de $|\Delta\theta_c|$ como função de θ_c para círculos de diferentes raios.

Por outro lado, de acordo com (6.7a), dado o círculo C de raio angular ρ a variação $\Delta\rho = \rho' - \rho$ é máxima quando a direção do centro de C corresponde ao eixo de $\vec{\beta}$, i.e. $\theta_c = 0$ ou $\theta_c = \pi$. Nestes casos, as posições detectadas dos centros dos círculos permanecerão fixas. O raio do círculo na direção do movimento, que corresponde a $\theta_c = 0$, diminui, enquanto o raio do círculo oposto ($\theta_c = \pi$) aumenta. A cota superior para $|\Delta\rho|$ é

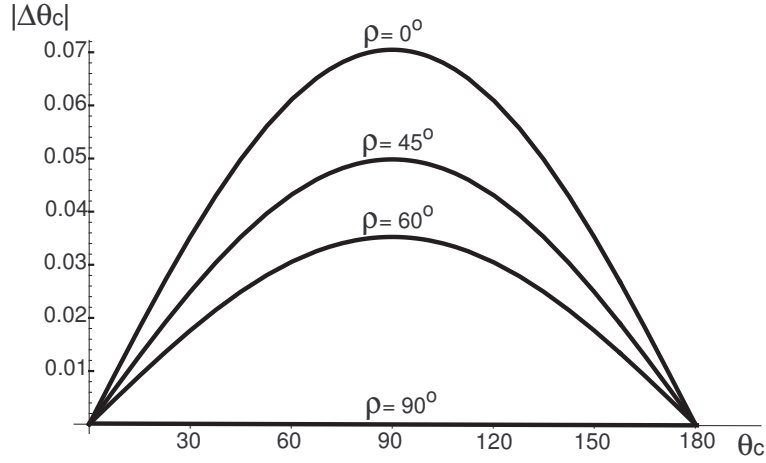


Figura 6.2: O valor da variação no raio dos círculos $\Delta\theta_c$ para círculos de diferentes raios e posição dos centros. Na figura, usamos o valor obtido a partir da magnitude do componente de dipolo na RCF [42], $\beta = 1,23 \times 10^{-3}$.

claramente $|\beta|$ novamente. A Figura 6.3 mostra o comportamento de $\Delta\rho$ como função de ρ para diferentes valores de θ_c . Claramente os círculos centrados no plano perpendicular ao eixo da velocidade $\vec{\beta}$ têm seus raios inalterados. A variação de ρ é máxima para círculos com centros ao longo do eixo β . Fica claro ainda que $\Delta\rho$ é uma função decrescente de ρ para $0^\circ \leq \theta_c < 90^\circ$, e crescente para $90^\circ < \theta_c \leq 180^\circ$.

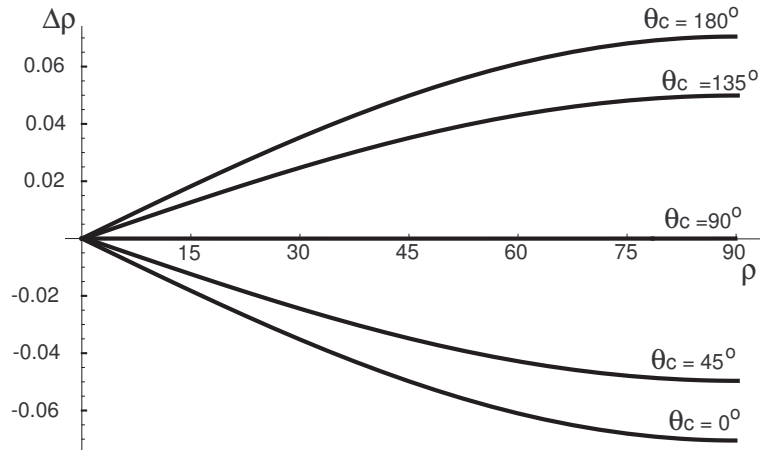


Figura 6.3: O valor da variação no raio dos círculos $\Delta\rho$ para círculos de diferentes raios e posição dos centros. Na figura, usamos novamente $\beta = 1,23 \times 10^{-3}$.

É fácil obter a partir de (6.7a) que no caso $\theta_c = 90^\circ$ o raio permanece inalterado. A cota superior para $|\Delta\rho|$ obviamente é $|\beta|$. Assim, dois círculos correlacionados terão no máximo uma diferença de raio de $2|\beta|$ quando detectados por um observador O' em movimento com velocidade $c\beta$ em relação ao referencial comóvel local.

Completando a caracterização completa dos efeitos de aberração da RCF devidos ao movimento próprio do observador, consideremos a variação da disposição angular dos

pontos ao longo dos círculos detectados pelos observadores O e O' . Para isso, definimos o círculo $C(\hat{q}_c, \rho)$ de centro $\hat{q}_c$ e raio ρ . Um ponto $\hat{q} = (\theta, \varphi)$ ao longo de C pode ser decomposto da seguinte forma:

$$\hat{q} = \vec{p} + \hat{q}_c \cos \rho, \quad (6.8)$$

onde $\vec{p}$ se localiza no plano do círculo, e é portanto ortogonal a $\hat{q}_c = (\theta_c, \varphi_c)$. O vetor $\vec{p}$ é o complemento ortogonal da projeção de $\hat{q}$ em $\hat{q}_c$, e o fator $\cos \rho$ é introduzido para garantir que $\hat{q}$ permaneça na esfera celeste. Tomamos como nosso ponto de referência a partir do qual medimos o ângulo de fase α o ponto $\vec{p}_0$, o complemento ortogonal do ponto superior do círculo (i.e., com o maior valor da coordenada θ), $\hat{q}_0 = (\theta_c - \rho, \varphi_c)$. Devido à simetria do problema, é evidente que φ_c permanece invariante quando o referencial é mudado. Escrevemos então

$$\cos \alpha = \frac{\vec{p} \cdot \vec{p}_0}{\sin^2 \rho}. \quad (6.9)$$

Claramente, $|\vec{p}| = |\vec{p}_0| = \sin \rho$. Com um pouco de álgebra, obtemos então

$$\cos \alpha' = \cos \alpha + \left(\frac{\beta}{1 + \beta \cos \theta} \right) \sin \rho \sin \theta_c \sin^2 \alpha, \quad (6.10)$$

onde $\cos \theta = \cos \rho \cos \theta_c + \sin \rho \sin \theta_c \cos \alpha$.

A diferença $\alpha' - \alpha$ pode ser obtida a partir da Equação (6.10), e é de ordem β . Podemos portanto afirmar que, até primeira ordem em β , $\cos \alpha' = \cos \alpha + (\alpha' - \alpha) \sin \alpha$. Expandindo o restante da expressão também até a primeira ordem em β obtemos então o deslocamento angular explicitamente

$$\alpha' = \alpha - \beta \sin \rho \sin \theta_c \sin \alpha. \quad (6.11)$$

É claro, a partir de (6.11), que para um dado círculo de raio ρ o módulo do deslocamento angular $|\Delta\alpha|$ é máximo para $\theta_c = \pi/2$ e $\alpha = \pm\pi/2$. A cota superior para $|\Delta\alpha|$ é claramente $|\beta|$, e corresponde à configuração onde $\theta_c = \pi/2$, $\rho = \pi/2$ e $\alpha = \pi/2$ simultaneamente. Novamente, dois círculos antipodais terão uma defasagem de no máximo $2|\beta|$. A figura 6.4 mostra a variação de $|\Delta\alpha|$ como função de α para diferentes valores do raio e posição dos centros dos círculos. Note ainda que para círculos maiores (maior ρ) a variação de fase é maior.

Combinando as análises acima, a figura 6.5 ilustra a combinação dos efeitos sobre o raio, posição do centro e distorção angular em um círculo no céu genérico devido à mudança de referencial. Na figura, pontos distribuídos uniformemente ao longo do círculo $C(\hat{q}_c, \rho)$ detectado pelo observador comóvel O serão observados pelo observador O' (coincidente com O no instante $t = t_0$, mas em movimento em relação ao referencial comóvel)

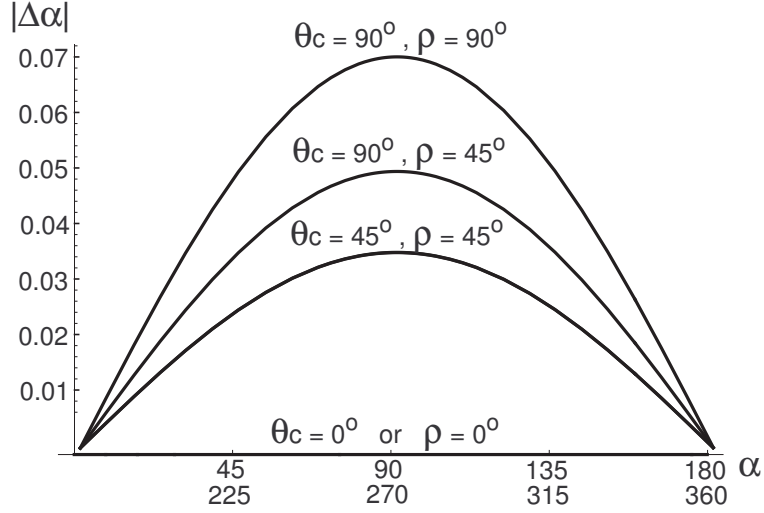


Figura 6.4: O valor do deslocamento angular $|\Delta\alpha|$ para círculos de diferentes raios e centros. Na figura, usamos $\beta = 1,23 \times 10^{-3}$.

ainda ao longo do círculo transformado $C'(\tilde{q}'_c, \rho')$, mas a sua distribuição será concentrada na direção do movimento.

6.3 Consequências para buscas futuras

Em resumo, mostramos acima que os círculos no céu detectados por um observador O *também serão detectados como círculos por um observador O'* em movimento relativo a e posição coincidente com O . Mas os círculos detectados por O e O' terão em geral diferentes raios, centros e distribuições de anisotropias ao longo do seu perímetro. Pares de círculos identificados por isometrias, que são idênticos para um observador comóvel, serão em geral diferentes para observadores em movimento em relação ao referencial comóvel.

Provamos ainda que para qualquer círculo no céu a variação máxima no valor do raio, posição do centro ou fase, é $\pm\beta$. Para observações efetuadas a partir da nossa galáxia, sabemos a partir do valor medido para a amplitude do momento de dipolo da radiação cósmica de fundo que estamos nos movendo em relação ao referencial comóvel local com velocidade $\beta = 1,23 \times 10^{-3}$, [42] o que implica em uma variação máxima para cada uma destas quantidades de $\simeq \pm 0,07^\circ = 4,2'$. Assim, círculos antipodais cujo centro se encontra na direção de $\hat{\beta}$ terão raios cuja diferença é no máximo $0,14^\circ$. Este valor é muito pequeno, abaixo da resolução máxima do WMAP ($0,25^\circ = 15'$) [158], mas próxima da resolução esperada para a missão Planck (pelo menos $0,16^\circ = 10'$), a ser lançada neste ano [162], e possivelmente acima da resolução de eventuais futuras missões.

O movimento da Terra em torno do Sol, e o movimento de um satélite em torno da

Terra, têm direção variável, e correspondem a aproximadamente $\beta_{Terra} = 7 \times 10^{-6}$ e $\beta_{Satelite} = 2 \times 10^{-6}$. Estes valores são irrisórios quando comparados ao movimento da Via Láctea como um todo (que tomamos como constante nas escalas de tempo em que as medidas são efetuadas e teses são escritas), e portanto não precisam ser considerados no presente contexto.

Poder-se-ia argumentar que mesmo se este efeito de aberração for significativo, seria mais fácil simplesmente inverter as transformações (6.1) e (6.2) para obter o resultado correspondente a um observador comóvel. De fato, para fontes discretas este é um procedimento relativamente simples [109]. Mas os mapas da RCF são intrincadas combinações de medições feitas em várias frequências, por um longo período de tempo, e cobrindo regiões diferentes do céu com diferentes intensidades. É um procedimento não trivial combinar estes dados de forma a não introduzir erros sistemáticos (c.f. Capítulo 2).

Assim, a aplicação da desaberração nos dados brutos provavelmente seria computacionalmente intensa, e precisaria renormalizar a intensidade da cobertura das diferentes regiões do céu. Por outro lado, a alteração pura e simples dos mapas pixelizados resultantes iria criar pixels de tamanhos diferentes, distribuídos de forma inhomogênea, ou exigiria uma nova pixelização que de alguma maneira combinasse pixels vizinhos e os redividisse. O procedimento descrito neste capítulo permite calcular a partir de qual resolução a busca por círculos no céu precisaria lançar mão de mapas de RCF desaberrados, ou então laboriosamente corrigir os parâmetros de cada círculo antes de procurar por pares correlacionados.

Finalmente, vale a pena enfatizar que os efeitos aqui descritos que resultam da aberração da RCF descrevem completamente os efeitos puramente geométricos devidos ao movimento próprio do observador (o que na prática corresponde ao movimento próprio da nossa galáxia como um todo). No caso de observações de círculos no céu, a aberração pode ser elegantemente decomposta em três efeitos que alteram as propriedades dos círculos. Lembramos ainda que estes efeitos são fenômenos puramente locais, e não dependem nem da curvatura do espaço-tempo ao longo das trajetórias dos fótons incidentes, nem da taxa de expansão do Universo.

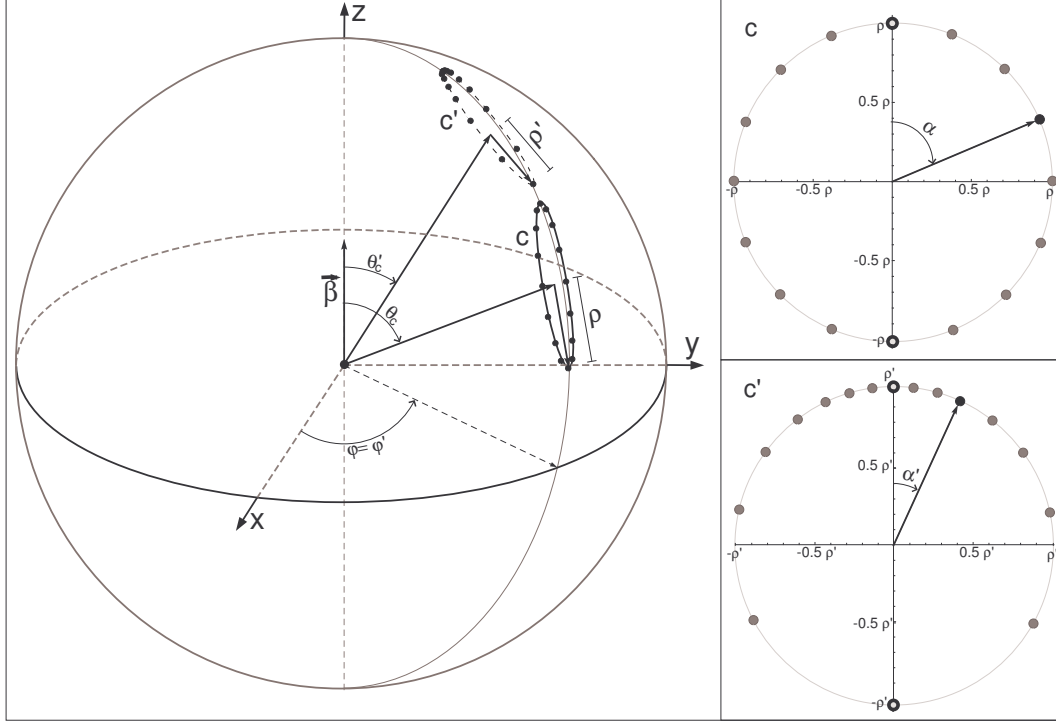


Figura 6.5: No painel esquerdo mostramos uma ilustração de todos os efeitos causados pela mudança de referencial na aparência dos círculos no céu: O círculo sólido $C(\hat{q}_c, \rho)$, detectado pelo observador O , é detectado pelo observador O' como o círculo tracejado $C'(\hat{q}'_c, \rho')$. Note a mudança na posição do centro, a variação do raio, e o efeito de deslocamento de fase ao longo do perímetro dos círculos. À direita, o efeito de deslocamento da fase é mostrado em maiores detalhes. Pontos uniformemente distribuídos em $C(\hat{q}_c, \rho)$ (painel superior, à direita) são deslocados ao longo do perímetro de $C'(\hat{q}'_c, \rho')$, se concentrando na direção do movimento (painel inferior). Os pontos situados ao longo do meridiano definido situando a origem das coordenadas na direção do movimento ($\alpha=0$ e $\alpha = 180^\circ$) não são afetados por este efeito. Na figura, mostramos o efeito de aberração real correspondente a $\beta = 0,8$.

Considerações finais

*Este não é o fim. Não é nem mesmo o começo do fim. É,
talvez, o fim do começo.*

Winston S. Churchill

Os paradigmas físicos bem-sucedidos contêm em seu cerne a semente de sua própria superação. À medida que o objeto de estudo vai se tornando mais bem compreendido, e que os parâmetros livres das teorias usadas para entendê-lo são determinados com crescente precisão pelas observações, as discrepâncias e fatos inexplicados remanescentes se tornam mais salientes, até serem por sua vez explicados por um entendimento mais completo da natureza. Não há por que supor que o atual paradigma cosmológico relativístico, que inclui um Big Bang quente seguido por uma fase inflacionária, será diferente. Apesar de seu enorme sucesso, sabemos que este entendimento é incompleto, e que em particular a topologia cósmica ainda está por ser determinada. Esta tese não teve a pretensão de determinar tal topologia, mas ela se insere em uma linha de pesquisa que tem feito avanços notáveis no sentido de procurar por sinais de uma topologia não trivial com reais chances de sucesso. Esperamos que nossa contribuição, embora modesta, seja útil neste contexto para avançar o entendimento de como e sob quais condições a topologia poderia ser detectada, e quais seriam as consequências observacionais se isto acontecesse.

A presença de uma significativa anisotropia estatística em mapas de radiação cósmica de fundo não tem ainda uma explicação comumente aceita, mas se inclui entre as discrepâncias entre teoria e observação que talvez tenham consequências importantes para o nosso entendimento do Universo. Os indicadores direcionais de anisotropia que desenvolvemos foram usados aqui para verificar que a anisotropia em grandes escalas detectada é estatisticamente significativa, robusta, e concentrada em torno de uma direção específica. Mas além disso, estes indicadores são ferramentas, que podem ser aplicadas a diversos outros problemas. No contexto do estudo da topologia cósmica, duas aplicações vêm

a mente. Em primeiro lugar, seria interessante verificar a hipótese segundo a qual a anisotropia detectada teria origem topológica; em particular, talvez seja possível determinar assinaturas topológicas para diferentes variedades multiplamente conexas. De forma análoga, se a anisotropia em grandes escalas que detectamos tiver origem em algum tipo de contaminação de foreground, então seria importante determinar o efeito desta contaminação na eficácia dos métodos de detecção da topologia cósmica que fazem uso de mapas de radiação cósmica de fundo, tais como o método de círculos no céu.

Na medida em que pode ser um empecilho para a detecção da topologia cósmica, a contaminação de foreground se assemelha aos efeitos do movimento próprio de nossa galáxia em relação ao referencial comóvel local. Fomos capazes de caracterizar completamente as conseqüências deste último efeito para a detecção de círculos no céu; mas existem ainda vários outros fatores que podem complicar ou comprometer a detectabilidade da topologia, cujos efeitos ainda precisam ser melhor compreendidos.

Dada a ausência da determinação efetiva da topologia cósmica, uma das poucas coisas que podemos afirmar a seu respeito são as condições de detectabilidade das diversas variedades possíveis, em termos dos parâmetros dos modelos para o conteúdo material do Universo. Vimos como fazer isto sistematicamente para o modelo Λ CDM e gás de Chaplygin generalizado e, em particular, notamos que a detectabilidade de uma variedade depende de forma crucial da escolha do modelo. Caso a topologia das seções espaciais seja determinada, porém, seria possível estudar o problema inverso da detectabilidade, e usar as distâncias entre imagens múltiplas do mesmo ponto ou objeto como um conjunto de observações astrofísicas capaz de restringir, conjuntamente com outros observáveis, os parâmetros do modelo cosmológico.

Por tudo que já dissemos acima, a procura por sinais de uma topologia não trivial é uma tarefa árdua, mas os dividendos de uma busca bem sucedida seriam consideráveis. Para aumentar as chances de sucesso, uma estratégia de busca eficiente se torna necessária. Como vimos, a forma local de variedades não-planas detectáveis no chamado limite inflacionário consiste basicamente em isometrias translacionais ou quase translacionais, o que simplifica bastante as buscas. O estudo da forma local tem o potencial de produzir outros resultados de aplicação geral, tais como as cotas independentes da posição para o comprimento da menor geodésica que obtivemos para espaços-lente não globalmente homogêneos.

Os anos recentes viram surgir a chamada cosmologia de precisão, o sucesso do modelo cosmológico padrão, e talvez até mesmo as primeiras evidências de uma nova física pós-relativística. São tempos interessantes de se viver quando se é um cosmólogo, e repletos de oportunidades. Em particular, no que se refere ao estudo da topologia cósmica, em-

bora um progresso significativo tenha ocorrido recentemente, ainda há muito a ser feito. Nos sentimos honrados pela oportunidade de dar nossa pequena contribuição para esta empreitada.

Apêndice A

O espectro de potência angular e a física das anisotropias

Embora seja somente uma entre várias formas disponíveis atualmente para estudar a cosmologia, a análise das anisotropias da RCF tem sido provavelmente a mais proveitosa, em particular no que se refere ao Universo primordial. As razões para isto são claras. Com o surgimento de observatórios orbitais, a tecnologia disponível torna possível medições extremamente precisas da radiação fóssil de microondas em todo o céu. Esta radiação é um retrato bastante fiel da época da recombinação; contaminações posteriores têm amplitudes relativamente pequenas, e em geral são facilmente identificáveis por sua natureza não-térmica. Finalmente, a emissão da RCF, e a evolução de suas anisotropias, derivam de processos físicos bem entendidos agindo em um regime linear. Tudo isto permitiu aos cosmólogos confirmar as linhas mestras do modelo padrão da cosmologia, e ao mesmo tempo procurar por indícios de uma nova física que vá além deste paradigma.

Uma das razões por que o estudo da RCF é tão proveitoso é que a física do Universo antes, durante e depois da época em que se deu o desacoplamento entre matéria e radiação é bem entendida. Após o período no qual se acredita que houve inflação, e antes da reionização, o entendimento corrente dos processos de termodinâmica relativística (o livro de Tolman é uma boa introdução ao tema [170]) tem se mostrado robusto, e razoavelmente consensual, pelo menos desde a década de 70 (exceto pelo ajuste de parâmetros livres). Os recentes avanços da cosmologia de precisão refinaram a nossa visão a respeito da composição material do Universo¹, mas confirmaram a teoria subjacente [81] (veja a

¹Em particular, eles sugerem fortemente que o Universo é (quase) plano, e composto primordialmente

Figura A.1).

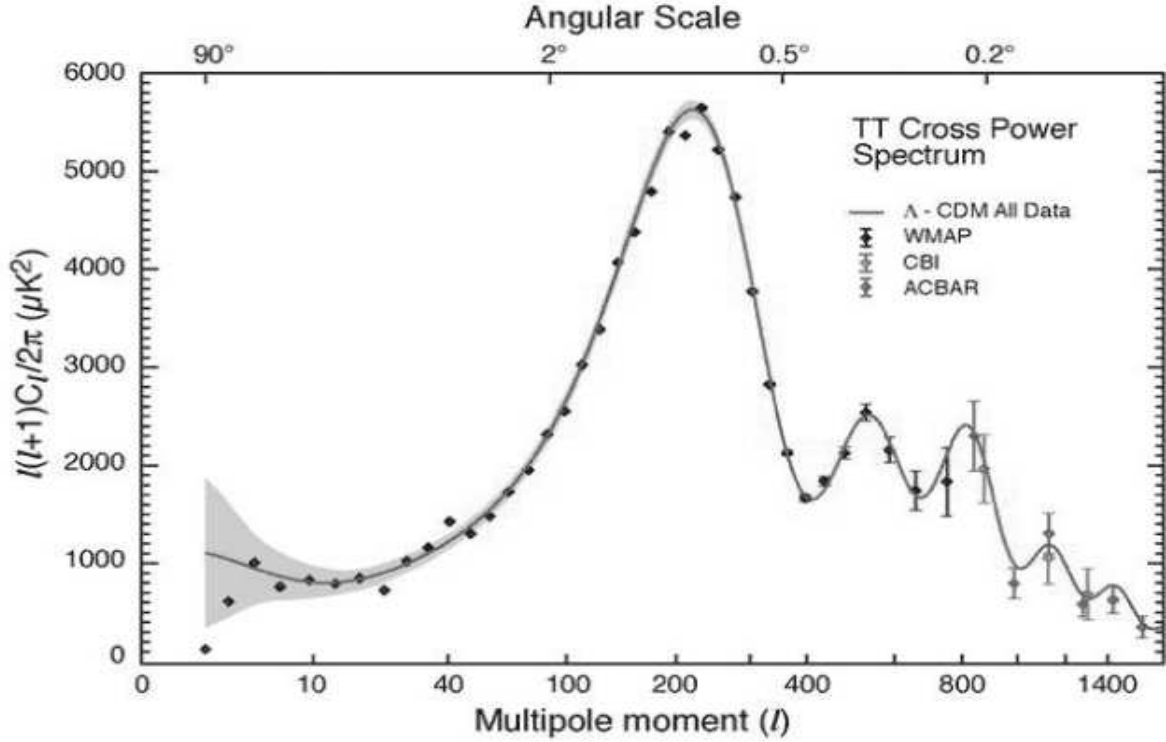


Figura A.1: O espectro de potência angular da radiação cósmica de fundo obtido pelo WMAP e outros experimentos. A linha sólida é um modelo teórico Λ CDM usando os melhores parâmetros disponíveis. Note o baixo valor do quadrupolo e octopolo. *Fonte: D. Miller, (2007).*

São cerca de oito² os parâmetros livres mais importantes do modelo padrão da cosmologia (o valor varia dependendo do número e tipo de componentes materiais que são levados em conta) [150]. Supõe-se que o Universo é bem descrito por uma métrica FLRW em grandes escalas (i.e., é aproximadamente homogêneo e isotrópico), com pequenas perturbações primordiais que deram origem às estruturas de aglomerados e superaglomerados de galáxias que observamos atualmente. Em princípio são dois os tipos de perturbações lineares possíveis; as chamadas adiabáticas, ou isoentrópicas (i.e., onde não há transferência de calor; diferenças de temperatura se devem unicamente à variações de densidade); e as perturbações de isocurvatura, nas quais a densidade de matéria-energia (e portanto a curvatura) permanece sempre constante. Os dados atuais aparentemente descartam

de matéria escura fria e energia escura, embora a determinação da natureza exata destes componentes ou sua detecção direta permaneça ainda em aberto.

²Em geral: H_0 , a constante de Hubble (ou a sua versão adimensional, $h = H_0/100 \text{ Mpc s km}^{-1}$); os parâmetros que determinam a lei de potência para a escala das perturbações primordiais A e n ; a razão entre a magnitude de perturbações escalares e tensoriais, r ; a profundidade ótica do espalhamento Thomson devido à reionização, τ ; e os parâmetros de densidade.

a possibilidade de que perturbações de isocurvatura tenham um papel significativo [41]. Acredita-se que ondas gravitacionais primordiais também tenham deixado sua marca na RCF na forma de perturbações tensoriais, mas o seu efeito é pequeno em relação às perturbações escalares.

As perturbações de densidade geradas por modelos inflacionários são normalmente quase invariantes por escala³. Devido à sua pequena amplitude em relação à média, a evolução destas perturbações primordiais pode ser decomposta em modos de Fourier independentes, cada um evoluindo independentemente. Cada modo é sujeito principalmente a duas ações opostas: a gravidade, que tende a reforçar as sobredensidades, e a pressão de radiação, que tende a desfazê-las. A combinação das duas forças induz oscilações harmônicas, que alternam sobre- e sub-densidades (a frequência de oscilação é determinada pela velocidade do som no meio), até ocorrer o desacoplamento entre matéria e radiação.

Após o desacoplamento, a força restauradora devida à pressão de radiação desaparece, e flutuações de densidade da matéria iniciam um (inicialmente lento) processo de colapso gravitacional, que continua até hoje, e deu origem à estrutura em grandes escalas do Universo atual. A radiação, após ser emitida pela última vez, se propaga pelo Universo subitamente transparente, sua temperatura determinada pela temperatura intrínseca da matéria, e pelos efeitos Doppler devidos ao movimento radial induzido pelas oscilações e ao potencial gravitacional das flutuações de densidade. Como o desacoplamento é simultâneo em todo o Universo, a RCF detectada por um observador comóvel atualmente terá sido emitida por uma esfera (com espessura correspondente à duração do processo de desacoplamento), que convencionalmente é denominada superfície de último espalhamento (SUE).

A evolução das perturbações lineares de densidade, e em particular a sua dependência dos valores dos parâmetros livres, é bem entendida. A relação entre perturbações de densidade e anisotropias de temperatura da RCF, cuja informação essencial está contida em seu espectro de potência, também é conhecida. Assim, o detalhamento do espectro de potência angular da RCF, além de verificar a validade geral do modelo cosmológico padrão, permite restringir com precisão inédita diversos destes parâmetros.

³Esta terminologia pode dar origem a confusão. Normalmente, um processo estocástico é dito invariante por escala se o seu espectro não possui uma escala de frequência característica, o que implica em uma lei de potência $P(\lambda \vec{k}) = \lambda^n P(\vec{k})$, onde $P(\vec{k})$ é a amplitude associada ao modo $\vec{k}$. Em cosmologia, por outro lado, o termo invariante por escala denota perturbações em que não só não há uma frequência característica, como a amplitude é a mesma para todos os modos no momento em que cada um deles cruza (i.e., se torna menor que) o horizonte de Hubble, o que corresponde no exemplo acima a $n = 1$.

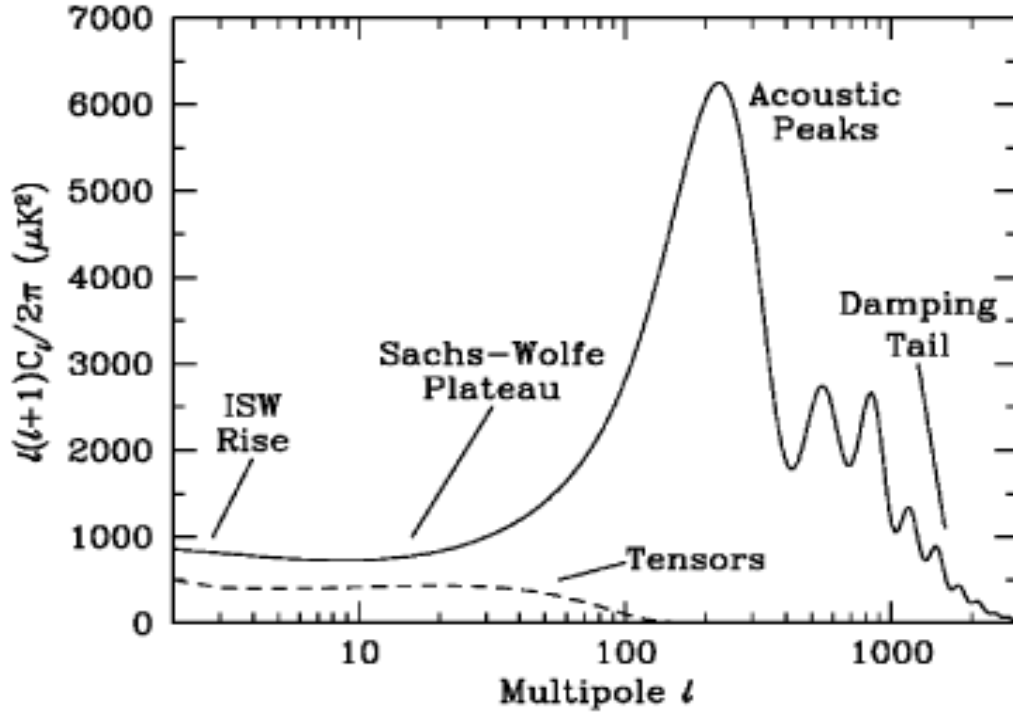


Figura A.2: O espectro de potência angular da radiação cósmica de fundo calculado teoricamente a partir de um modelo Λ CDM usando os melhores parâmetros disponíveis. As principais características mencionadas no texto são indicadas na figura. *Fonte: D. Scott & G.F. Smoot (2006).*

Processos físicos diferentes agem em escalas angulares diferentes. Como pode ser visto na Figura A.2, é possível identificar três regimes principais no espectro de potência angular:

$\ell \lesssim 100$ (o platô Sachs-Wolf)

Estes componentes correspondem aos modos com escalas maiores que o raio de Hubble, e que portanto não tiveram tempo até o desacoplamento de evoluir significativamente. Preservam portanto o aspecto original das flutuações de densidade. Neste regime, as flutuações de temperatura são proporcionais às flutuações do potencial gravitacional com contribuições com sinais opostos de flutuações intrínsecas de temperatura e desvio para o vermelho gravitacional (o efeito Sachs-Wolfe, veja [184]). Se este espectro das flutuações de densidade é invariante por escala, então a quantidade $\ell(\ell + 1)C_\ell$ é aproximadamente constante, o famoso platô. Outros efeitos, como a variação ao longo do tempo dos potenciais gravitacionais (o chamado efeito Sachs-Wolfe integrado, [110]), e modos tensoriais, ocasionam um suave aumento do valor dos C_ℓ para $\lesssim 10$. Para ℓ s baixos, porém, a incerteza induzida pela variância cósmica é bastante significativa, e é difícil fazer afirmações

categóricas. Os experimentos recentes confirmaram a presença do platô, embora os valores para o quadrupolo ($\ell = 2$) e octopolo ($\ell = 3$) sejam inusitadamente baixos em relação aos valores esperados. São nestas escalas que uma possível topologia não trivial para as seções espaciais do Universo possa ter um efeito mensurável.

$100 \lesssim \ell \lesssim 1000$ (os picos e vales acústicos)

Esta é possivelmente a parte mais informativa do espectro de potência. É caracterizada por uma sucessão de picos e vales bem definidos, consequência direta das oscilações acústicas. Cada modo de Fourier tinha uma frequência característica de oscilação. Os modos cuja frequência era tal que passavam exatamente por um máximo ou mínimo de densidade quando se deu a recombinação deram então contribuições desproporcionadamente significativas para as flutuações de temperatura quando a RCF foi emitida. Assim, os multipolos associados às escalas angulares correspondentes aos comprimentos de onda destes modos de Fourier ressonantes com o tempo de recombinação são máximos no espectro de potência. Os picos ímpares se devem a máximos de sobredensidade, e os pares, a máximos de subdensidades. Particularmente notável é o primeiro pico. Ele corresponde ao modo que teve tempo para evoluir exatamente $1/4$ do seu período de oscilação até a recombinação, e portanto com o gradiente de densidade indo de zero até o máximo. É possível, a partir de primeiros princípios, deduzir o comprimento de onda deste modo, assim como calcular o raio da SUE. Mas, usando somente geometria, se sabemos o tamanho de um objeto distante, e o ângulo que ele subtende em um espaço de curvatura constante, é possível relacionar sua distância com a curvatura do espaço. Isto pode ser inferido diretamente da parte espacial da métrica de FLRW (1.10). Integrando separadamente ds ao longo de um raio de comprimento r (integrando $d\chi$ entre 0 e r/a e mantendo ϕ e θ constantes), e em seguida em torno de um círculo definido por este raio (integrando $d\phi$ entre 0 e 2π , com $\chi = r/a$, $\theta = \pi/2$), é fácil ver que perímetro e raio estão relacionados por $P = 2\pi a f(r/a)$, onde $f(\chi) = \chi$, $\sin \chi$ e $\sinh \chi$ para $k = 0, 1$ e -1 respectivamente. Assim, o valor exato de ℓ onde se situa o primeiro pico acústico determina com considerável precisão a curvatura (com ℓ maiores para curvaturas negativas, e menores para curvaturas positivas), e portanto o parâmetro $h^2\Omega_0$. O cálculo exato é mais elaborado, porque tanto o raio da SUE quanto o comprimento de onda do primeiro modo ressonante dependem (embora mais fracamente) de outros parâmetros, mas pode ser encontrado em [37]. Alguns outros efeitos mais sutis também precisam ser levados em conta (e.g., o espalhamento Thomson de fótons da RCF após a reionização induz um fator de atenuação para ℓ 's mais altos, alterando ligeiramente a posição dos picos).

A posição e amplitude dos picos também restringem significativamente outros parâmetros livres da teoria (notadamente a fração de bárions), mas uma explicação detalhada está fora do escopo da presente revisão. Veja [81] para maiores detalhes e para uma noção intuitiva da relação entre os picos e os vários componentes materiais. Os valores para os parâmetros cosmológicos obtidos a partir da análise dos picos acústicos detectados pelo WMAP, e a comparação dos resultados observacionais com modelos teóricos, estão disponíveis em [158] e [159].

$\ell \gtrsim 1000$ (a cauda de atenuação)

O desacoplamento entre matéria e radiação não se deu instantaneamente. À medida que a recombinação foi ocorrendo, cada vez menos partículas carregadas estavam presentes para interagir com a radiação. Como consequência deste enfraquecimento gradual do acoplamento, a radiação se difundiu entre regiões de sobre- e subdensidades, diminuindo os gradientes de pressão e amortecendo as oscilações. Claramente, quanto menor a escala das oscilações, mais rápido o amortecimento. Este efeito, chamado de amortecimento Silk, é importante para multipolos cujas escalas angulares correspondem à modos com comprimento de onda menor do que a espessura da SUE, e vai se tornando mais importante até praticamente suprimir os multipolos com $\ell \gtrsim 2000$.

Apêndice B

A classificação das 3–variedades

Certo dia, enquanto pescava em um rio, Nasrudin viu um viajante se aproximar, exausto, da margem oposta. "Como posso chegar ao outro lado do rio?"- gritou o viajante. Nasrudin respondeu: "Você já está do outro lado do rio!"

Anônimo, antigo conto sufi

Apresentamos aqui uma breve classificação das 3–variedades multiplamente conexas, para cada um dos espaços de cobertura possíveis $\widetilde{M}$: $\mathbb{R}^3$, $\mathbb{S}^3$ ou $\mathbb{H}^3$. Para uma abordagem acessível do assunto por dois grandes matemáticos que trabalharam nesta área, veja [168]. Para um estudo sistemático, leia [189]. Vamos nos limitar aqui às variedades orientáveis, que têm maior interesse físico [88]. Para desenvolver a intuição a respeito destas variedades, existe um interessante programa de computador desenvolvido por Jeffrey Weeks, 'Curved Spaces' [178], que permite visualização do seu domínio fundamental.

B.1 $\widetilde{M} = \mathbb{R}^3$ (curvatura $K = 0$)

O grupo de isometrias de $\mathbb{R}^3$ pode ser gerado simplesmente pelo produto das translações ao longo, e das rotações em torno, de 3 eixos linearmente independentes em $\mathbb{R}^3$. Podemos escrever portanto $I(\mathbb{R}^3) = \mathbb{R}^3 \times SO(3)$. Toda isometria livre e propriamente descontínua em $I(\mathbb{R}^3)$ é combinação de uma rotação com uma translação não trivial ao longo do eixo de rotação. Chamaremos de **isometria helicoidal** aquelas nas quais o eixo de translação coincide com o eixo de rotação.

Todos os espaços quocientes compactos de $\mathbb{R}^3$ tem um espaço de recobrimento toroidal.

O mesmo vale para qualquer dimensão. Por exemplo, na garrafa de Klein, como vimos, o gerador não-translacional, se aplicado duas vezes, gera uma translação. Esta translação, combinada com um segundo gerador translacional, gera um subgrupo toroidal.

A classificação dos grupos de holonomia de $\mathbb{R}^3$ é bem conhecida há mais de um século, tendo sido publicada pela primeira vez pelo matemático e cristalografista russo Yevgraf Fyodorov, em 1885 (a classificação é apresentada e discutida em [143] no contexto da topologia cósmica). Temos:

Variedades abertas (não-compactas)

- $\mathcal{E}$: O próprio espaço de cobertura $\mathbb{R}^3$;
- $\mathcal{J}_\theta$: O produto de um cilindro por $\mathbb{R}^2$, com $\Gamma_{\mathcal{J}_\theta}$ gerado por uma isometria helicoidal com uma fase θ (o valor do componente de rotação das isometrias 'tipo parafuso'), que é um parâmetro livre;
- $\mathcal{T}_1$: O produto de um 2-toro por $\mathbb{R}$, com $\Gamma_{\mathcal{T}_1}$ gerado por duas translações;
- $\mathcal{K}_1$: O produto da garrafa de Klein por $\mathbb{R}$, com $\Gamma_{\mathcal{K}_1}$ gerado por uma translação e uma isometria helicoidal, perpendicular à translação e de fase π .

Variedades compactas

- E_1 (3-toro): Γ_{E_1} gerado por 3 translações. O domínio fundamental formado por um paralelepípedo; única 3-variedade plana compacta globalmente homogênea;
- E_2 : Γ_{E_2} gerado por 2 translações, perpendiculares ao eixo de rotação de uma isometria helicoidal de fase π . O domínio fundamental é um paralelepípedo de base reta, hexagonal ou retangular;
- E_3 : Γ_{E_3} gerado por 2 translações ortogonais entre si, e perpendiculares a uma isometria helicoidal de fase $\pi/2$. O domínio fundamental é um paralelepípedo de base reta e quadrada;
- E_4 : Γ_{E_4} gerado por 2 translações em direções que perfazem um ângulo de $\pi/3$ entre si, perpendiculares a uma isometria de fase $2\pi/3$. O domínio fundamental é um prisma reto de base hexagonal com lados opostos de igual tamanho, e as faces laterais identificadas pelas 2 translações e pela isometria definida pela diferença e entre elas, juntamente com suas respectivas inversas;

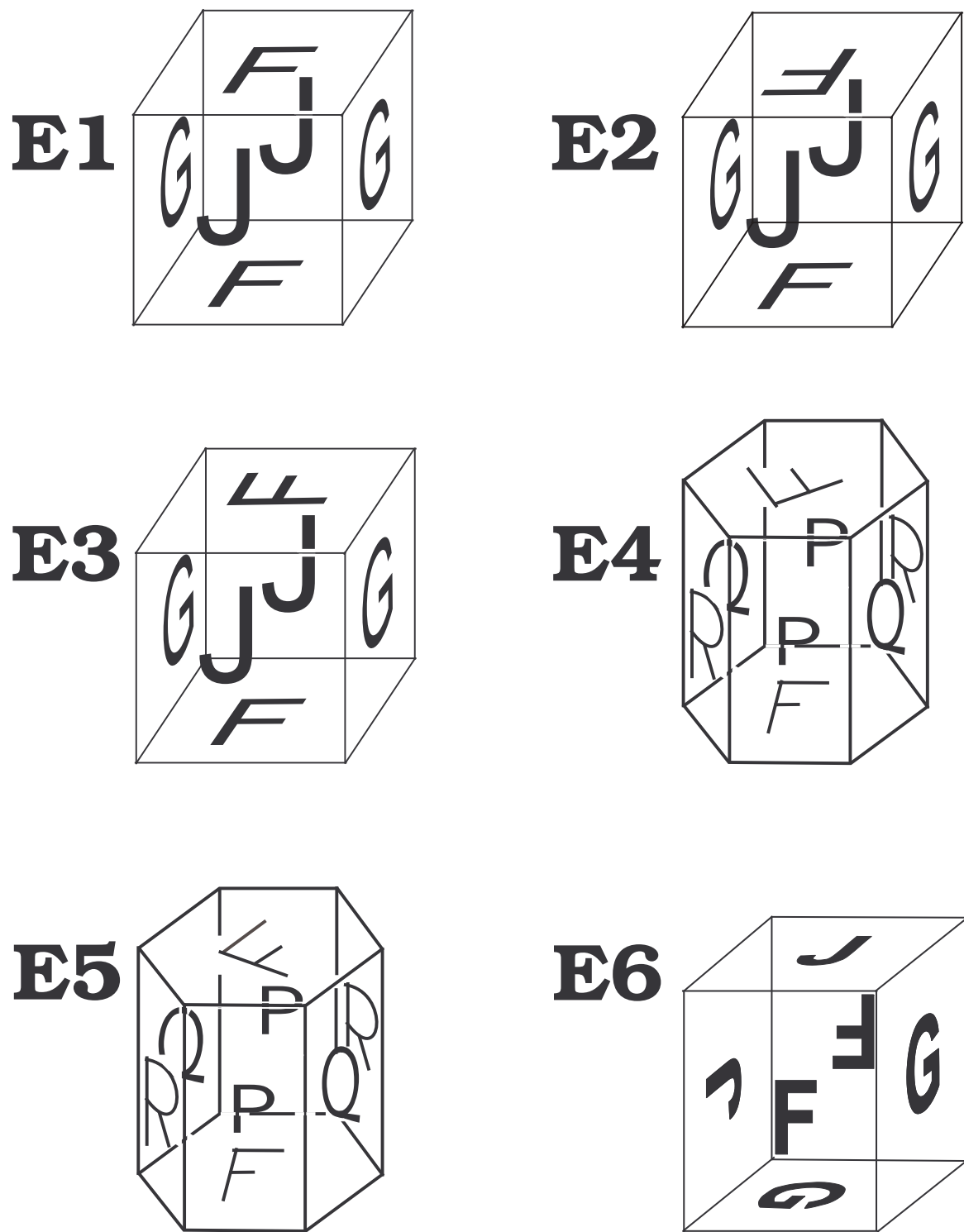


Figura B.1: Domínios fundamentais das 3–variedades euclidianas compactas e orientáveis. A identificação entre as faces é feita segundo indicado pela orientação das letras. Em todos os casos, os parâmetros livres das holonomias em questão foram fixados de forma a obter a representação mais simples. Para uma descrição dos geradores de Γ veja o texto. *Fonte: A. Teixeira, modificado pelo autor.*

- E_5 : Γ_{E_5} gerado por 2 translações perfazendo um ângulo de $\pi/3$ entre si, perpendiculares a uma isometria de fase $\pi/3$. O domínio fundamental é um prisma reto de base hexagonal, com as faces laterais identificadas pelas 2 translações e pela isometria definida pela diferença e entre elas, juntamente com suas respectivas inversas;
- E_6 (Hantzche-Wendt): Γ_{E_6} gerado por 3 isometrias não-helicoidais de fase π (os eixos de rotação são ortogonais entre si, e perfazem um ângulo de $\pi/4$ com os eixos de translação, também ortogonais entre si). O domínio fundamental é um dodecaedro rombóide. No caso em que os componentes translacionais de todos os três geradores são de igual comprimento, o domínio fundamental se reduz a um cubo.

As variedades compactas descritas acima são ilustradas na Figura B.1.

B.2 $\widetilde{M} = \mathbb{S}^3$ ($K > 0$)

O grupo de isometrias $I(\mathbb{S}^3)$ é isomorfo a $SO(4)$, o grupo das rotações¹ em $\mathbb{R}^4$. Obviamente, como a variedade $\mathbb{S}^3$ é compacta, ela é coberta por um número finito de cópias do domínio fundamental (que é a ordem do grupo de holonomias).

A 3-esfera pode ser representada como o conjunto de pontos $x = \{x_1, x_2, x_3, x_4\}$ em $\mathbb{R}^4$ tal que $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$, e suas operações de simetria podem ser expressas em termos de multiplicação de quatérnions (uma extensão não-comutativa para quatro dimensões do conceito de números complexos). Brevemente, podemos representar cada ponto como um quatérnion unitário $q(x) = x_1 + x_2i + x_3j + x_4k$, e a multiplicação se dá segundo as regras $i^2 = j^2 = k^2 = -ikj = ijk = -jik = jki = -1$. É possível mostrar que toda isometria $\gamma \in I(\mathbb{S}^3)$ pode ser escrita como $rq(x)s = \gamma x$, onde r e s são quatérnions unitários. Esta forma é unívoca, a menos da troca simultânea dos sinais de r e s . Assim, existe um isomorfismo $SO(4) \approx \mathbb{S}^3 \times \mathbb{S}^3 / \{\pm(1, 1)\}$, onde cada ponto de $\mathbb{S}^3$ corresponde a um quatérnion dotado do produto definido acima. Cada subgrupo finitos de holonomias de $SO(4)$ pode ser obtido a partir de um par de subgrupos finitos de quatérnions, digamos Φ_d e Φ_e , que agem, no todo ou em parte, de forma multiplicativa respectivamente à direita e à esquerda de cada ponto em $\mathbb{S}^3$. Os subgrupos finitos de $SO(4)$ podem ainda ser classificadas em:

i) Ação simples, quando $\Phi_d = \{1\}$ ou $\Phi_e = \{1\}$

¹Em esferas de dimensão par, todas as rotações têm pontos fixos, e $SO(n+1)$ não admite ação livre. Assim, e.g., a 2-esfera não admite variedade quociente além dela mesma, e não existe nenhuma 2-variedade de curvatura constante multiplamente conexa.

As isometrias correspondentes a ação simples à direita ou à esquerda geram translações de Clifford (como é fácil verificar calculando diretamente $|xq - q|$). Todas as variedades desta classe, listadas abaixo, são portanto globalmente homogêneas, e têm uma correspondência um a um com os subgrupos finitos de $\mathbb{S}^3$, entendido como o grupo dos quatérnions unitários:

- Os grupos cíclicos Z_n , de ordem n , gerados por uma translação de $2\pi/n$; $n = 1$ é a própria $\mathbb{S}^3$ e $n = 2$ é o espaço projetivo RP^3 , que corresponde à identificação entre duas metades da esfera por uma inversão através da origem²;
- os grupos binários diédricos D_m^* , de ordem $2m$, $m \geq 2$, gerados por uma translação de $2\pi/m$ e uma inversão através da origem;
- o grupo binário tetraédrico T^* , de ordem 24 ;
- o grupo binário octaédrico O^* , de ordem 48 ;
- o grupo binário icosaédrico I^* , de ordem 120; $\mathbb{S}^3/I^*$ tem como domínio fundamental o famoso dodecaedro de Poincaré.

Os últimos três grupos são chamados grupos binários poliédricos, pois têm origem em uma cobertura dupla dos grupos de simetrias dos poliedros regulares (lembrando que o cubo e o dodecaedro têm as mesmas simetrias que respectivamente o octaedro e o icosaedro), que são os únicos subgrupos finitos das rotações em $\mathbb{R}^3$, $SO(3)$.

ii) Ação dupla, quando todas as permutações de elementos de Φ_d e Φ_e são usadas.

iii) Ação conjunta, quando nem todas as permutações de elementos de Φ_d e Φ_e são usadas.

Estas duas últimas classes são geradas pela ação de (partes de) grupos translacionais, à direita e à esquerda. Para uma classificação completa, veja [57].

É possível mostrar que em geral, para uma escolha apropriada de eixos, toda isometria $\gamma \in SO(4)$ propriamente descontínua pode ser representada como uma ação sobre um par de números complexos, (z_1, z_2) , $z_1 = x_1 + x_2i$, $z_2 = x_3 + x_4i$, da seguinte maneira:

$$\gamma(u, v) = (e^{2\pi i \frac{n}{p}} z_1, e^{2\pi i \frac{nq}{p}} z_2) . \quad (\text{B.1})$$

Tais isometrias são as geradoras únicas dos chamados espaços-lente, denotados $L(p, q)$, que dependendo dos valores de q e p podem ter ação simples, dupla ou conjunta. Visualmente, seu nome deriva da forma lenticular do seu domínio fundamental. Os grupos

²Note que neste caso, o gerador do grupo γ é tal que (em notação quaterniônica) $\gamma q = -q$. Assim, γ comuta com todos os elementos de $SO(4)$, e portanto RP^3 é não só globalmente homogêneo, mas também globalmente isotrópico. Esta é a única 3-variedade multiplamente conexa com esta propriedade.

cíclicos Z_n correspondem ao caso especial $q = 1$, quando o espaço é globalmente homogêneo.

É possível também provar que 3-variedades esféricas são **rígidas**, isto é, que se duas variedades tem grupos fundamentais isomorfos, elas serão necessariamente isométricas (a menos de homotetia). Note que o mesmo não acontece em variedades planas.

B.3 $\widetilde{M} = \mathbb{H}^3$ ($K < 0$)

Existem dois fatos notáveis a respeito das isometrias $I(\mathbb{H}^3)$ [169]: Não existe uma classificação completa de seus subgrupos discretos de ação livre e propriamente descontínua, e nenhuma delas é uma translação de Clifford. O domínio fundamental sempre depende da posição do ponto-base. De fato, as variedades hiperbólicas são de trato muito mais difícil que as demais. Lembramos que toda superfície orientável com curvatura constante positiva ou nula é respectivamente uma 2-esfera ou um 2-toro, mas são infinitos os tipos de superfícies com curvatura constante negativa (uma para cada genus de toro acima de 2), com poliedros fundamentais com número de lados arbitrariamente grandes. Da mesma maneira, em três dimensões, enquanto o número de faces do poliedro fundamental não passa de 12 no caso plano e 120 no caso esférico, variedades hiperbólicas podem ter poliedros fundamentais com um número arbitrário de faces. Tendo em vista tudo isso, não é possível estudar as 3-variedades hiperbólicas por meio de uma classificação completa em termos de famílias, e existem somente alguns teoremas de aplicação geral como guia. Notadamente, G. D. Mostow mostrou em 1968 [112] que variedades hiperbólicas fechadas (i.e., completas e compactas) e orientáveis, com dimensão maior que 2, são rígidas.

Atualmente, a menor variedade deste tipo conhecida é a chamada variedade de Weeks [180], para a qual o volume é $V = 0,9247 a^3$ (a é o raio de curvatura, que é igual ao fator de escala para a métrica FLRW)³. Sabe-se, graças a Agol e Dinfield, que para qualquer variedade hiperbólica fechada orientável $V > 0,67 a^3$. [2]

O grupo de isometrias de $\mathbb{H}^3$ é isomorfo ao grupo de transformações fracionárias agindo sobre os complexos, $PSL(2, \mathbb{C})$, i.e., transformações $\gamma(z) = z'$ tais que

$$z' = \frac{az + b}{cz + d}, \quad a, b, c, d \in \mathbb{C}, \quad ad - bc = 1. \quad (\text{B.2})$$

Uma categoria de variedades hiperbólicas de volume finito, mas não compactas, consiste nas chamadas variedades ditas cuspidais. A sua não-compacidade se deve a presença

³Em um artigo ainda não publicado [56], David Gabai et al. afirmam ter provado que a variedade de Weeks é a variedade hiperbólica fechada e orientável de menor volume possível.

de uma cúspide que vai se afinilando até um ponto singular no infinito. Tais variedades surgem como o limite de seqüências de variedades compactas de volume crescente [169], e são particularmente úteis como ponto de partida para gerar as demais variedades compactas, através da chamado cirurgia de Dehn hiperbólica, ou mais propriamente preenchimento de Dehn, que consiste em 'colar' um toro sólido a uma variedade cuspidal, o que resulta em uma nova variedade hiperbólica possivelmente compacta. A forma como o toro e a variedade original são identificados é indexada por dois inteiros co-primos n_1 e n_2 denominados winding numbers (não há um termos correspondente em português de uso corrente, e por isto usaremos o termo em inglês). Assim, por exemplo, a variedade hiperbólica de Weeks é indicada por $m003(-3, 1)$, onde $m003$ é a variedade cuspidal (e, incidentalmente, a variedade cuspidal com menor volume) e -3 e 1 são os winding numbers. O infatigável Weeks desenvolveu também SnaPea [176], um programa de computador (disponível gratuitamente na internet), que permite estudar as variedades resultantes deste procedimento, e, juntamente com Hodgson [80], produziu um censo contendo 11031 variedades hiperbólicas compactas orientáveis, incluindo o seu volume e outros invariantes topológicos tais como o comprimento da menor geodésica fechada.

Apêndice C

Curvas de confiança e observações de supernovas

No campo das observações, o acaso favorece a mente preparada.

Louis Pasteur

C.1 Observações e regiões de confiança

Considere uma quantidade Y , medida observacionalmente. Como já repetimos diversas vezes, incertezas observacionais são inerentes ao processo de medida. Sob condições bastante gerais, podemos tratar Y como uma variável aleatória gaussiana, o que implica em que medições sucessivas Y retornarão valores y distribuídos da forma

$$\rho(y) = \frac{1}{\mathcal{N}} \exp \left[-\frac{1}{2} \frac{(y - y_0)^2}{\sigma_Y^2} \right], \quad (\text{C.1})$$

onde $\mathcal{N}$ é um fator de normalização, y_0 é o valor 'real' de Y (e média de Y por simetria), e σ_Y^2 é a sua variância (que pode ser obtida por exemplo pelo cálculo do desvio-padrão de uma série de medidas). Daqui para a frente, suporemos que todas as variáveis observacionais são gaussianas, salvo afirmativa em contrário. O procedimento a seguir pode ser repetido em princípio para outras distribuições, mas o tratamento caso a caso necessário neste caso pode ser bastante complicado.

Considere agora que Y seja uma função de um conjunto de outros observáveis, independentes entre si e também gaussianos: $Y = Y(\mathbf{X})$; $\mathbf{X} = (X^1, X^2 \dots X^N)$, cada um dos

quais também tem um desvio-padrão σ_{x^i} e média $\bar{X}^i$. Podemos propagar os erros para obter

$$Y(\mathbf{X}) \simeq Y(\bar{\mathbf{X}}) + \left. \frac{\partial Y}{\partial X^1} \right|_{X^1=\bar{X}^1} (X^1 - \bar{X}^1) + \dots + \left. \frac{\partial Y}{\partial X^N} \right|_{X^N=\bar{X}^N} (X^N - \bar{X}^N) . \quad (\text{C.2})$$

Escrito desta forma, fica claro que a variável aleatória gaussiana $Y(\mathbf{X})$ é a soma de variáveis aleatórias gaussianas $Y(\bar{X})$, $\left. \frac{\partial Y}{\partial X^1} \right|_{\bar{X}^1} (X^1 - \bar{X}^1)$, $\dots$, $\left. \frac{\partial Y}{\partial X^N} \right|_{\bar{X}^N} (X^N - \bar{X}^N)$, com médias $y_0, 0, \dots, 0$ e variâncias σ_Y^2 , $\left(\left. \frac{\partial Y}{\partial X^1} \right|_{\bar{X}^1} \sigma_{X^1} \right)^2$, $\dots$, $\left(\left. \frac{\partial Y}{\partial X^N} \right|_{\bar{X}^N} \sigma_{X^N} \right)^2$. Mas a soma de variáveis gaussianas é também sempre uma variável gaussiana, cuja média e variância são a soma das médias e variâncias das variáveis constituintes. Então, a medição Y retornará valores y com distribuição

$$\rho(y) = \frac{1}{\mathcal{N}} \exp \left[-\frac{1}{2} \frac{(y - y_0)^2}{\sigma_Y^2 + \sum_{j=1}^N \left(\left. \frac{\partial Y_i}{\partial X^j} \right|_{\bar{X}^j} \sigma_{X^j} \right)^2} \right] . \quad (\text{C.3})$$

Até agora, só levamos em conta uma única observação de Y e $\mathbf{X}$. Considere agora uma série de dados, que mede Y para diversos valores de $\mathbf{X}$; $(Y_1(\mathbf{X}), Y_2(\mathbf{X}), \dots, Y_K(\mathbf{X}))$. Este conjunto de medições retornará valores $(y_1, y_2, \dots, y_K)$ com uma distribuição conjunta $\rho(y_1, y_2, \dots, y_K)$. Mas se as observações forem independentes (e.g., medições de objetos distintos), então $\rho(y_1, y_2, \dots, y_K) = \rho(y_1)\rho(y_2)\dots\rho(y_K)$, e temos simplesmente

$$\rho(y_1, y_2, \dots, y_K) = \frac{1}{\mathcal{N}} \exp \left[-\frac{1}{2} \chi^2 \right] , \quad (\text{C.4})$$

onde $\mathcal{N}$ é a normalização da distribuição, obtida integrando ρ por todos os y_i , e $\chi^2 = \chi_y^2$ segue uma distribuição de χ -quadrado (c.f. (1.25) com K graus de liberdade, e é dada por

$$\chi_Y^2 = \sum_{i=1}^K \frac{(y_i - y_{0i})^2}{\sigma_i^2} , \quad (\text{C.5})$$

onde escrevemos $\sigma_i^2 = \sigma_{Y_i}^2 + \sum_{j=1}^N \left(\left. \frac{\partial Y_i}{\partial X^j} \right|_{\bar{X}^j} \sigma_{X^j} \right)^2$.

De fato, nada do procedimento acima exige que as observações independentes meçam a mesma variável. Podemos combinar diferentes séries de dados, $Y, Y^1, \dots, Y^{(r)}$ simplesmente escrevendo

$$\chi^2 = \chi_Y^2 + \chi_{Y^1}^2 + \dots + \chi_{Y^{(r)}}^2 . \quad (\text{C.6})$$

A distribuição de probabilidade $\rho(y_i; y_{0i})$ em (C.4) expressa simplesmente qual a probabilidade de um conjunto de observáveis Y_i com valores y_{0i} ser medido com valores y_i . Mas na prática observacional, o que se dá é exatamente o contrário. Temos um conjunto

de observações y_i , mas não sabemos a priori qual o valor 'real' das variáveis que estamos medindo. Mas note que (C.4) é simétrico em relação aos dois conjuntos de variáveis. Assim, podemos com igual rigor afirmar que a distribuição de probabilidade $\rho(y_i; y_{0i})$ expressa qual a probabilidade de um dado conjunto de observáveis ter valores reais y_{0i} se foi medido com valores y_i . Deste ponto de vista, cada escolha de y_{0i} corresponde a uma hipótese a ser testada estatisticamente. Neste caso, o valor mais provável de y_{0i} , ou melhor ajuste, é aquele que maximiza ρ (ou, de forma equivalente, minimiza χ^2 , ou ainda minimiza os quadrados das variáveis) e as regiões de confiança são superfícies de contorno do tipo $\chi^2 = \text{const.}$, que em termos destas variáveis são elipsóides com eixos $r\sigma_i$, com $r > 0$. O valor de confiança de tais superfícies de contorno, formalmente obtido integrando (C.5) por todo seu interior, tem valores bem conhecidos para múltiplos inteiros do desvio padrão, i.e., 68,2% para $r = 1$, 95,4% para $r = 2$, 99,6% para $r = 3$, etc.

Até agora, as espressões acima só são úteis para impor limites diretamente sobre valores dos y_{0i} , em termos dos quais as superfícies de contorno não dizem muito. Mas o caso que mais nos interessa é obter restrições sobre parâmetros que não são diretamente observáveis. Por isto, considere novamente a variável observável $Y(\mathbf{X}, \mathbf{W})$ (o caso para múltiplas variáveis é essencialmente idêntico, simplesmente substituímos χ^2 pela soma (C.4)), função agora não somente dos observáveis dependentes $\mathbf{X} = (X^1, X^2 \dots X^N)$, mas também das variáveis $\mathbf{W} = (W^1, W^2 \dots W^M)$, que não são diretamente observáveis, e que procuramos restringir. Neste caso, para um conjunto de K observações de Y e $\mathbf{X}$, y_i^{obs} e x_i^{obs} , temos um χ^2 que é claramente uma função apenas de $\mathbf{w}$,

$$\chi_Y^2(\mathbf{w}) = \sum_{i=1}^K \frac{(y_i^{obs} - y_i(\mathbf{x}_i^{obs}, \mathbf{w}))^2}{\sigma_i^2(\mathbf{x}_i^{obs}, \mathbf{w})}. \quad (\text{C.7})$$

Definimos assim o melhor ajuste de y_i^{teo} , como sendo $y_i^{teo}(\mathbf{w}_0) | \chi_Y^2(\mathbf{w}_0) = \min \chi_Y^2(\mathbf{w})$, e as superfícies de contorno $\chi_Y^2(\mathbf{w}) = \text{const.}$, no espaço paramétrico M -dimensional definido por $\mathbf{w} = (w^1, w^2, \dots, w^M)$, determinados pela distribuição de probabilidade $\rho(\mathbf{w}) = \frac{1}{N} \exp[\chi_Y^2(y(\mathbf{w}))]$ (ou $\frac{1}{N} \exp \left[\sum_{i=1}^K \chi_{Y_i}^2(y_i(\mathbf{w})) \right]$ no caso K observações independentes). Note que esta distribuição em geral não é gaussiana. Podemos marginalizar uma ou mais variáveis integrando a densidade de probabilidade por todos os seus valores possíveis; a distribuição resultante é a chamada probabilidade marginal das variáveis restantes. Se estamos interessados em $w^1 \dots w^r$, e supomos que as probabilidades marginais de $w^{r+1} \dots w^M$ são independentes, marginalizamos para obter ¹

¹Formalmente, pelo teorema de Bayes, temos $\rho(A|B)P(B) = \rho(A \cap B) = \rho(B|A)\rho(A)$, onde $\rho(A|B)$ é a probabilidade condicional de A dado B , $\rho(B)$ é a probabilidade marginal de B e $\rho(A \cap B)$ é a probabilidade conjunta de A e B . Integrando os dois lados em A , obtemos $P(B) = \int \rho(B|A)\rho(A)dA$.

$$\begin{aligned}\rho(w^1 \dots w^r) &= \int \rho(w^1 \dots w^r | w^{r+1} \dots w^M) \rho_p(w^{r+1}) \dots \rho_p(w^M) dw^{r+1} \dots dw^M, \\ \rho(w^1 \dots w^r) &= \int \frac{\rho(\mathbf{w}) \rho_p(w^{r+1}) \dots \rho_p(w^M)}{\int \rho(\mathbf{w}) dw^1 \dots dw^r} dw^{r+1} \dots dw^M.\end{aligned}\tag{C.8}$$

As distribuições $\rho_p(w^i)$ são chamados 'priors', podendo ser vínculos observacionais externos ou hipóteses consideradas razoáveis; no caso trivial de vínculos que não trazem informações adicionais à distribuição $\rho(\mathbf{w})$, temos $\rho_p(w^{r+1}) \dots \rho_p(w^M) = \int \rho(\mathbf{w}) dw^1 \dots dw^r$, e C.8 se simplifica: $\rho(w^1 \dots w^r) = \int \rho(\mathbf{w}) dw^{r+1} \dots dw^M$. Outro caso especial importante consiste nos priors gaussianos no limite $\sigma_i \rightarrow 0$, os chamados priors fortes, cuja distribuição toma a forma $\rho(w^i) = \delta(w^i - w_0^i)$ (e.g., o popular $\rho(\Omega_0) = \delta(\Omega_0 - 1)$).

C.2 A distância-luminosidade

Em astrofísica, medidas de fotometria são normalmente expressas em termos da magnitude, uma medida logarítmica invertida. Se dois objetos têm luminosidade ℓ_1 e ℓ_2 , então suas magnitudes aparentes m_1 e m_2 serão tais que

$$m_2 - m_1 = 2,5 \log_{10} \frac{\ell_1}{\ell_2}.\tag{C.9}$$

A magnitude de um objeto é uma quantidade adimensional, e portanto sem uma escala característica; é então necessário definir uma luminosidade padrão para que adquira um valor numérico independente (e não simplesmente relativo). Tradicionalmente, o brilho da estrela Vega era usado para definir a luminosidade correspondente à magnitude aparente 0 (atualmente o ponto zero é definido em termos de fluxo total para diferentes faixas de comprimento de onda). A magnitude aparente obviamente depende da distância entre fonte e observador; para uma medida de brilho intrínseco, a chamada magnitude absoluta, ou M , é definida como sendo a magnitude aparente de um objeto situado a 10 parsecs de distância. Para relacionarmos as magnitudes aparente e absoluta, é preciso determinar a relação entre distância e luminosidade.

Em um espaço plano, a luminosidade ℓ de um objeto distante cai com o quadrado da sua distância d ,

$$\ell = \frac{\Phi}{4\pi d^2},\tag{C.10}$$

onde Φ é o fluxo (em $erg\ s^{-1}$).

De forma mais geral, a luminosidade é simplesmente uma medida do fluxo de energia por unidade de área. Assim, no caso mais geral de um espaço de curvatura constante,

a luminosidade será inversamente proporcional à área de uma esfera que tem a fonte luminosa no centro e o observador em sua superfície. Uma esfera de raio $d = a\chi$ tem área $4\pi a^2 f^2(\chi)$, onde $f(\chi)$ é χ , $\sin \chi$ ou $\sinh \chi$ para respectivamente curvaturas 0, $1/a^2$ e $-1/a^2$ (como pode ser deduzido da parte espacial da métrica FLRW (1.10)). Além disso, se fonte e observador estiverem em movimento relativo, a luminosidade será afetada pelo efeito Doppler, por um fator de $(1+z)^{-2}$ (as variações de frequência e de intensidade da radiação contribuem cada uma com um fator de $1+z$). Assim, no caso de uma curvatura espacial constante, a luminosidade é dada por

$$\ell = \frac{\Phi}{4\pi(1+z)^2 a^2 f^2(\chi)} . \quad (\text{C.11})$$

Podemos recuperar a forma (C.10) para ℓ definindo a chamada 'distância-luminosidade', $d_L = (1+z)f(\chi)$. Reescrevendo (C.9) em termos de distâncias, temos então

$$m - M = 5 \log_{10} d_L - 5 . \quad (\text{C.12})$$

Se um objeto astronômico tem a sua magnitude absoluta conhecida, então a medição de sua magnitude aparente determina o valor de d_L , por meio de (C.12) (em um caso observacional idealizado, obviamente).

Em um Universo com métrica FLRW, χ pode ser escrito em termos do desvio para o vermelho da fonte (e dos parâmetros do conteúdo material) através da relação desvio para o vermelho—distância (3.7), e o raio de curvatura é simplesmente o inverso do quadrado do fator de escala, $1/a_0^2 = K = H_0^2 |1 - \Omega_0|^2$. Portanto,

$$d_L = \frac{1+z}{H_0 \sqrt{|1 - \Omega_0|}} f(\chi) . \quad (\text{C.13})$$

C.3 Supernovas SN Ia como velas-padrão

Supernovas do tipo SN Ia [76] são eventos altamente energéticos, que acredita-se ocorreram quando uma anã branca absorve matéria suficiente (em geral de uma companheira binária gigante vermelha) para que sua massa ultrapasse o chamado limite de Chandrasekhar, acima do qual a pressão degenerada dos elétrons é incapaz de impedir um colapso gravitacional. A explosão resultante tem duas grandes virtudes (do ponto de vista de um astrônomo): é extremamente luminosa, sendo portanto detectável a distâncias con-

sideráveis, e tem uma magnitude absoluta de pico aproximadamente constante, $M^{pico} \simeq -18.9^2$. Estas características tornam as SNs Ia candidatas ideais a velas-padrão, úteis para determinar a distância de galáxias com desvio para o vermelho cosmológico significativo, e assim mapear a função de Hubble, $H(z)$ (veja, e.g., [132] e [86]).

Na seção (4.2), por exemplo, usamos medições da curva de luz de 157 SNs Ia (a amostra *gold* de [144]), de onde se obtém as suas magnitudes aparente e absoluta, e portanto, a partir de (C.12), é possível tabelar $\log d_L^{obs}{}_i$ em função de z_i , e da incerteza associada a cada medida, $\sigma_{\log(d_L^{obs}{}_i)}^2$. Por (C.13), vemos ainda que a função qui-quadrado dos observáveis referentes às supernovas é uma função dos parâmetros cosmológicos Ω_m, Ω_Λ e H_0 , $\chi_{SN}^2 = \chi_{SN}^2(\Omega_m, \Omega_\Lambda, H_0)$. Reduzimos o espaço paramétrico à dimensão 2 marginalizando por H_0 , $\mathcal{L} = \frac{1}{N} \int e^{-\chi_{SN}^2(\Omega_m, \Omega_\Lambda, H_0)/2} dH_0$. Estes dados são combinados com a detecção de um par de círculos identificados por uma translação de Clifford de distância conhecida $2r_{inj}$ e raio ρ_0 , cuja medição tem desvio padrão σ_ρ . O vínculo gerado pela detecção dos círculos identificados nas faces do dodecaedro de Poincaré é introduzido adicionando um termo extra ao qui-quadrado χ_{SN}^2 , desta vez usando χ_{lss} como variável (a partir de (3.7) com $z_{lss} = 1089$). Usando (4.2) para expressar (3.8), temos $\chi^2 = \chi_{SN}^2 + \chi_{I*}^2$, com

$$\chi_{I*}^2 = \frac{\left[\arctan\left(\frac{\tan r_{inj}}{\cos \rho_0}\right) - \chi_{lss}(\Omega_m, \Omega_\Lambda, z_{lss}) \right]^2}{\left[\frac{\partial \chi_{lss}}{\partial \rho} \Big|_{\rho=\rho_0} \sigma_\rho \right]^2}; \quad (C.14)$$

a derivada $\frac{\partial \chi_{lss}}{\partial \rho}$ pode ser obtida diretamente a partir de (4.2), $\frac{\partial \chi_{lss}}{\partial \rho} = \frac{\cos r_{inj}}{\cos^2 r_{inj} \cos^2 \rho + \sin^2 \rho}$.

Em (4.2), por outro lado, tomamos uma amostra com 253 SNs Ia compiladas a partir de [172] e [9]. Os valores tabelados no caso são z_i , $\log(D_L^{obs}{}_i)$ e a variância $\sigma_{\log(D_L^{obs}{}_i)}^2$ para cada supernova, onde $D_L = d_L H_0$ (que elimina a dependência explícita de H_0 em (C.13)). Descartamos supernovas com $z < 0,01$, cujo efeito Doppler devido ao movimento próprio é comparável ao efeito cosmológico, assim como aquelas cujo brilho sofre contaminação local significativa ($A_V > 0,5$). O conjunto restante consiste de 194 SNs Ia com $z_{\max} = 1,75$. Levamos em conta no qui-quadrado ainda a dispersão em z causada por velocidades próprias radiais v_i com distribuição gaussiana e $\sigma_v = 500 \text{ km/s}$. Propagando esta incerteza em σ_z , escrevemos

$$\chi_{SN}^2 = \sum_{i=1}^{194} \frac{\left[\log(D_L^{obs}{}_i) - \log(D_L(z_i, \chi(z_i))) \right]^2}{\sigma_{\log(D_L^{obs}{}_i)}^2 + \left(\frac{\partial \log D_L}{\partial z} \Big|_{z_i} \sigma_z \right)^2}. \quad (C.15)$$

Na distribuição resultante, as linhas de contorno são dadas em termos dos parâmetros de $\chi(z_i)$, em um modelo do tipo gás de Chaplygin generalizado.

²Certos parâmetros na curva de luz da supernova podem ser usados para determinar M^{pico} com precisão ainda maior.

Apêndice D

Teoria dos números aplicada

A matemática é a rainha das ciências, e a teoria de números é a rainha da matemática.

Carl Frederick Gauss

À matemática pura, e que ela nunca seja útil a alguém.

Godfrey Harold Hardy

A epígrafe acima é um antigo brinde entre os matemáticos puros; mas mesmo no século 19, quando foi proferido pela primeira vez, felizmente já não era levada muito a sério. Porém, entre todas as áreas da matemática, uma se sobressaía por seu apelo ao que tudo indicava puramente estético, sem nenhuma aplicação substancial aparente nas ciências ou engenharias.

A teoria dos números é de fato belíssima, mas os seus objetos de estudo, os números inteiros e a sua estrutura em termos de seus fatores primos, sempre pareceram particularmente alheios a qualquer aplicabilidade ao mundo físico. Ao longo dos anos, porém, ela foi aplicada a problemas concretos, em particular na área de criptografia. Mas não deixa de ser surpreendente que ela encontre aplicação na cosmologia, como vimos no Capítulo 5. Abaixo apresentamos sucintamente os elementos fundamentais da teoria dos números, assim como a prova dos lemas e teoremas que usamos no texto. São resultados que, embora relativamente simples, tem conseqüências potencialmente amplas para a detecção de um possível topologia não trivial para o universo.

D.1 Alguns resultados da teoria de números

D.1.1 Operações em módulo, e congruências

A maneira natural de estudar a estrutura aritmética de um número natural é através do exame do resto resultante da sua divisão por outros números naturais. Dizemos portanto que inteiros a e b são congruentes módulo m ($m > 0$) se existe algum inteiro r tal que $a = b + rm$. Neste caso, escrevemos $a \equiv b \pmod{m}$. A congruência define uma relação de equivalência associativa e distributiva entre os inteiros, que são divididos em m classes de equivalência.

Usaremos a notação (a, b) para denotar o máximo denominador comum entre a e b . Se $(a, b) = 1$, a e b são ditos 'primos entre si'. As seguintes propriedades valem

Se $ca \equiv cb \pmod{m}$, e $(c, m) = 1$, então $a \equiv b \pmod{m}$.

Se $a \equiv 0 \pmod{m}$, então $a/(a, m) \equiv 0 \pmod{p/(a, m)}$

Uma introdução legível à teoria dos números pode ser encontrada em [121]. Para uma abordagem mais completa sobre o tópico de frações continuadas, de onde vêm os principais resultados apresentados abaixo, veja [84].

O foco do que se segue é mantido unicamente na aplicação dos resultados ao estudo da topologia cósmica, e não há nenhuma pretensão de fornecer uma introdução completa ao tema; mas no que se refere aos tópicos abordados, procuramos criar uma apresentação auto-contida, embora forçosamente sucinta.

D.1.2 O teorema de Euler, e o pequeno teorema de Fermat

Um teorema historicamente importante para a teoria de números, atribuído a Euler, é provado abaixo e usado logo adiante. Fazemos uso aqui da chamada função μ de Möbius. Definimos $\mu(n)$, $n \in \mathbb{N}$ como a cardinalidade do conjunto de naturais menores que n que não possuem fatores em comum com n . Formalmente, $\mu(n) = \#\{r < n \mid (r, n) = 1\}$.

Teorema D.1 *Se n é um inteiro positivo e a é um inteiro tal que $(a, n) = 1$, então*

$$a^{\phi(n)} \equiv 1 \pmod{n}$$

.

Prova: Considere os conjuntos a) $r_1, r_2, \dots, r_{\phi(n)}$ com $r_i < n$, $(r_i, n) = 1$; e b) $ar_1, ar_2, \dots, ar_{\phi(n)}$.

Os elementos do conjunto $r_1, r_2, \dots, r_{\phi(n)}$ com $r_i < n$, $(r_i, n) = 1$ são dois a dois incongruentes módulo n (i.e., se $r_i \equiv r_j \pmod{n} \rightarrow i = j$), pois são todos menores que n . Da

mesma maneira, os elementos $ar_1, ar_2, \dots, ar_{\phi(n)}$ também são dois a dois incongruentes módulo n , já que como $(a, n) = 1$ temos que $ar_i \equiv ar_j \pmod{n} \rightarrow r_i \equiv r_j \pmod{n} \rightarrow i = j$. Além disso, pela mesma razão $(ar_i, n) = 1$, o que implica que $ar_i \equiv r_j \pmod{n}$ para algum j . Assim, cada ar_i é congruente a exatamente um r_j . Como ambos os conjuntos têm a mesma cardinalidade, podemos então escrever

$$ar_1 ar_2 \dots ar_{\phi(n)} \equiv r_1 r_2 \dots r_{\phi(n)} \pmod{n},$$

ou seja

$$a^{\phi(n)} r_1 r_2 \dots r_{\phi(n)} \equiv r_1 r_2 \dots r_{\phi(n)} \pmod{n}.$$

Cancelando os termos em comum (o que é possível já que $(r_i, n) = 1$) obtemos $a^{\phi(n)} \equiv 1 \pmod{n}$ ■

Note que se n é primo, $\phi(n) = n - 1$; neste caso o teorema se reduz ao chamado pequeno teorema de Fermat.

D.1.3 Frações continuadas

Qualquer número racional $\frac{q}{p}$, com q e p primos entre si (i.e., $(q, p) = 1$), pode ser escrito como uma fração continuada finita

$$\frac{q}{p} = a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \dots + \frac{1}{a_{k-1} + \frac{1}{a_k}}}}, \quad (\text{D.1})$$

onde $a_1 \in \mathbb{Z}$, $a_i \in \mathbb{N} \mid i > 1$ são os chamados quocientes parciais, a parte inteira de sucessivas aproximações da fração. Usaremos a representação $\frac{q}{p} \equiv [a_1, a_2, a_3, \dots, a_k]$, que é única a menos da equivalência $[a_1, \dots, a_n] = [a_1, \dots, a_n - 1, 1]$ se $a_n > 1$. Para restaurar a unicidade, vamos sempre tomar o último quociente parcial tal que $a_k > 1$. Note ainda que se $0 < q/p < 1$ então $a_1 = 0$, e que $p/q = 0 + 1/(q/p) = [0, a_1, \dots, a_k]$. Isto garante

Lema D.1 $p_{k-1} < \frac{p}{2}$

Prova: Lembrando que convencionamos $a_k \geq 2$, e que por construção $p_{k-2} > 0$, e portanto $p = a_k p_{k-1} + p_{k-2} > 2p_{k-1}$ ■

Fazemos uso desta notação para escrever ainda $p/q = [a_1, \dots, a_k] = [a_1, \dots, \alpha_j]$, onde $\alpha_j = [a_j, \dots, a_k]$.

Considere a seqüência de aproximações de q/p , denominadas convergentes:

$$\begin{aligned}
c_1 &= \frac{q_1}{p_1} = a_1 = [a_1] , \\
c_2 &= \frac{q_2}{p_2} = a_1 + \frac{1}{a_2} = [a_1, a_2] , \\
c_3 &= \frac{q_3}{p_3} = a_1 + \frac{1}{a_2 + \frac{1}{a_3}} = [a_1, a_2, a_3] , \\
&\vdots \\
c_k &= \frac{q}{p} = a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{\ddots \frac{1}{a_{k-1} + \frac{1}{a_k}}}}} = [a_1, a_2, a_3, \dots, a_k] ;
\end{aligned} \tag{D.2}$$

Os termos q_n e p_n podem ser obtidos recursivamente

$$q_i = a_i q_{i-1} + q_{i-2}; \quad q_0 = 0, \quad q_{-1} = 1 , \tag{D.3a}$$

$$p_i = a_i p_{i-1} + p_{i-2}; \quad p_0 = 1, \quad p_{-1} = 0 ; \tag{D.3b}$$

obviamente $p_k = p$ e $q_k = q$.

Usando as fórmulas acima recursivamente, provamos que

Lema D.2 $p_i q_{i-1} - p_{i-1} q_i = (-1)^i$ para todo $0 \leq i \leq k$

Prova: As fórmulas acima garantem trivialmente que

$$p_1 q_{-1} - p_{-1} q_1 = 1^0 = 1$$

então escrevemos

$$p_i q_{i-1} - p_{i-1} q_i = (a_i p_{i-1} + p_{i-2}) q_{i-1} - p_{i-1} (a_i q_{i-1} + q_{i-2})$$

$$p_i q_{i-1} - p_{i-1} q_i = p_{i-2} q_{i-1} - p_{i-1} q_{i-2}$$

Se $p_{i-1} q_{i-2} - p_{i-2} q_{i-1} = (-1)^{i-1}$ então a fórmula acima se reduz a $p_i q_{i-1} - p_{i-1} q_i = (-1)^i$.

Como ela é válida para $i = 0$, então por indução ela é válida para todo $i \geq 0$ ■

É possível provar usando as definições que $(q_i, p_i) = 1$, e também que

$$q_1 < \dots < q_i < \dots < q_k , \tag{D.4a}$$

$$p_1 < \dots < p_i < \dots < p_k . \tag{D.4b}$$

Podemos estender o resultado acima, notando que como $p/q = [a_1, \dots, \alpha_j]$, podemos usar exatamente o mesmo processo indutivo para escrever $p = \alpha_j p_{j-1} + p_{j-2}$ e $q = \alpha_j q_{j-1} + q_{j-2}$. Provamos assim um lema deveras útil

Lema D.3 $\frac{q}{p} - \frac{q_i}{p_i} = \frac{(-1)^{i+1}}{p_i(p_{i+1} + p_i/\alpha_{i+2})}$ para todo $0 \leq i < k$

Prova: Escrevemos

$$\begin{aligned} \frac{q}{p} - \frac{q_j}{p_j} &= \frac{\alpha_{j+1}q_j + q_{j-1}}{\alpha_{j+1}p_j + p_{j-1}} - \frac{q_j}{p_j} = \frac{p_j(\alpha_{j+1}q_j + q_{j-1}) - q_j(\alpha_{j+1}p_j + p_{j-1})}{p_j(\alpha_{j+1}p_j + p_{j-1})} \\ \frac{q}{p} - \frac{q_j}{p_j} &= \frac{p_jq_{j-1} - q_jp_{j-1}}{p_j(\alpha_{j+1}p_j + p_{j-1})} = \frac{(-1)^{j+1}}{p_j(p_{j+1} + p_j/\alpha_{j+2})} \end{aligned}$$

Onde definimos $\alpha_{k+1} \rightarrow \infty$ de modo que a fórmula acima valha para todo $0 \leq i < k$ ■

Usamos o resultado acima para expressar a quantidade $q^{\phi(p)-1}$ de uma forma mais simples

Lema D.4 $p_{j-1} \equiv (-1)^j q_j^{\phi(p_j)-1} \pmod{p_j}$; em particular $p_{k-1} \equiv (-1)^k q^{\phi(p)-1} \pmod{p}$

Prova: A partir do Lema D.3, podemos escrever $p_jq_{j-1} - p_{j-1}q_j = (-1)^j$, e portanto $p_{j-1}q_j \equiv (-1)^k \pmod{p_j}$. Por outro lado, o teorema de Euler (D.1) garante que $q^{\phi(p_j)-1}q_j \equiv 1 \pmod{p_j}$. Como por hipótese $(q_j, p_j) = 1$, podemos multiplicar os dois lados por $q^{\phi(p_j)-1}$ e obter $p_{j-1} \equiv (-1)^j q^{\phi(p_j)-1} \pmod{p_j}$ ■

Um resultado análogo vale para a fração p/q , cujos convergentes são sempre os inversos dos convergentes de q/p de mesma ordem: $p_{j-1} \equiv (-1)^j p_j^{\phi(q_j)-1} \pmod{p_j}$. Estes dois resultados são úteis para computar os pares p_i e q_i recursivamente a partir de p e q de uma forma simples (i.e., mais fácil de ser implementada computacionalmente).

Notamos de passagem que se tomássemos $a_k = 1$, então teríamos $p_{k-2} = p - p_{k-1} < p/2$; neste caso $p_{k-1} > p/2$, mas também $q^{\phi(p)-1} \equiv (-1)^k p_{k-1} \equiv (-1)^{k-1} p_{k-2} \pmod{p}$.

Um resultado de Lagrange demonstra o que intuitivamente já seria de se esperar: Os convergentes $\frac{q_n}{p_n}$ são de uma certa maneira as melhores aproximações possíveis de $\frac{p}{q}$.

Teorema D.2 (Lagrange) Se existirem racionais a/b e q/p com $p > b \geq 1$ e $(a, b) = (p, q) = 1$ tais que

$$\left| \frac{q}{p} - \frac{a}{b} \right| < \frac{1}{2b^2}$$

então $a = q_i$ e $b = p_i$ para algum i , ou seja, a/b é um convergente de p/q .

Prova: Existe sempre n tal que $p_n \leq b < p_{n+1}$. O teorema é provado trivialmente se $\frac{a}{b} = \frac{q}{p}$; caso contrário, suponha por absurdo que a desigualdade acima é verdadeira mas a/b não é um convergente de p/q . Então pelo teorema acima,

$$\left| p_n \frac{q}{p} - q_n \right| < \left| b \frac{q}{p} - a \right| < \frac{1}{2b} ,$$

e portanto

$$\left| \frac{q}{p} - \frac{q_n}{p_n} \right| < \frac{1}{2bp_n} \leq \frac{|bq_n - ap_n|}{2bp_n} = \frac{1}{2} \left| \frac{q_n}{p_n} - \frac{a}{b} \right| ;$$

mas

$$\left| \frac{q_n}{p_n} - \frac{a}{b} \right| = \left| \frac{q_n}{p_n} - \frac{q}{p} + \frac{q}{p} - \frac{a}{b} \right| \leq \left| \frac{q}{p} - \frac{q_n}{p_n} \right| + \left| \frac{q}{p} - \frac{a}{b} \right| ;$$

Substituindo as desigualdades, obtemos

$$\frac{1}{bp_n} \leq \left| \frac{q}{p} - \frac{q_n}{p_n} \right| + \left| \frac{q}{p} - \frac{a}{b} \right| \leq \frac{1}{2bp_n} + \frac{1}{2b^2}$$

Isto implica em que $b < p_n$, o que é uma contradição, provando o teorema ■

Terminamos esta seção com um pequeno resultado que garante haver um limite para o número máximo de convergentes que uma dada fração pode ter

Lema D.5 $k \leq 2 \log_2 2q$

Prova: Como $p_i = a_i p_{k-1} + p_{i-2}$ e também $p_i > p_{i-1}$ e $p_1 = 1$, temos que $p_2 \geq 2p_1$ e $p_i > 2p_{i-2}$. Portanto, $p = p_k > 2^{\frac{k+1}{2}}$. Rearranjando os termos, obtemos $k \leq 2 \log_2 p = 2 \log_2 p$. Podemos ainda melhorar esta cota lembrando que $\frac{q}{p} = 0 + \frac{1}{p/q}$. Aplicando o mesmo raciocínio à fração continuada de $\frac{p}{q}$, que tem $k' = k - 1$ componentes, verificamos que $k' \leq 2 \log_2 q$, e portanto $k \leq 2 \log_2 2q$. Note que já que por definição $q < \frac{p}{2}$, então $2 \log_2 2q < 2 \log_2 p$ ■

D.2 Aplicação a espaços-lente

D.2.1 Grupos finitos de simetrias de $\mathbb{S}^3$ e os espaços-lente

O grupo de isometrias da 3-esfera $\mathbb{S}^3$ é isomorfo à $SO(4)$, o grupo das rotações em $\mathbb{R}^4$. Uma holonomia esférica genérica pode ser entendida como automorfismo (sem pontos fixos) de $\mathbb{S}^3$ nela mesma. Genericamente, pode ser representada como a ação multiplicativa pela esquerda e pela direita de quaternions unitários a e b sobre um ponto representado pelo quaternion x .

$$\gamma_{ab}(x) = axb ; a, b, x \in \mathbb{H}, |a| = |b| = |x|^2 = 1 . \quad (\text{D.4c})$$

A classificação de todos os subgrupos finitos de holonomias de $\mathbb{S}^3$, cada um dos quais corresponde a uma forma espacial esférica possível, obtida pela primeira por Threlfall e Seifert [167], pode ser expressa em termos desta ação multiplicativa de grupos finitos de quaternions. Veja também a Seção B para uma discussão dos subgrupos assim gerados.

Quando a multiplicação se dá por um grupo de quatérnions somente à esquerda ou à direita (ação simples), as isometrias geradas são sempre translações de Clifford; quando a multiplicação é pela esquerda e pela direita, com todas as permutações dos elementos de dois grupos (ação dupla), existem sempre dois subgrupos translacionais (gerados pelas permutações dos grupos de quatérnions de um lado com a identidade do outro). Quando a multiplicação só ocorre para pares selecionados dos dois grupos (ação conjunta), translações de Clifford podem ou não estar presentes.

Uma discussão completa desta classificação, aplicada à topologia cósmica, pode ser encontrada em [57].

Uma outra representação possível das holonomias esféricas se dá em termos de uma ação cíclica sobre um par ordenado de complexos (u, v)

$$\gamma_{(p,q)}^n(z_1, z_2) = (e^{2\pi i \frac{n}{p}} z_1, e^{2\pi i \frac{nq}{p}} z_2) ; z_1, z_2 \in \mathbb{C}, |z_1|^2 + |z_2|^2 = 1 , \quad (\text{D.4d})$$

com $(p, q) = 1$, $n < p$. Para os nossos propósitos, tal representação é mais prática, e será usada daqui em diante. Por conveniência, denotaremos uma tal isometria como $\gamma_{(p,q)}^n$. Uma forma espacial esférica cujo grupo de holonomias é gerado por uma única isometria é chamado espaço-lente. Podemos provar que

Lema D.6 *O grupo das holonomias geradas por $\gamma_{(p',q')}^n$ também é gerado por $\gamma_{(p,q)}$ para alguma escolha de p e q .*

Prova: Escrevemos $n = n' \times (n, p')$, $p' = p \times (n, p')$ e $q' = q \times (n, p')$, e cancelamos os fatores (n, p') comuns. Então, pelo teorema de Euler, $n'^{\phi(p)} \equiv 1 \pmod{p}$. Assim, aplicando $\gamma_{(p',q')}^n$ $n'^{\phi(p)-1}$ vezes, obtemos $\gamma_{(p,q)}^{n'\phi(p)} = \gamma_{(p,q)}$; como os grupos gerados por $\gamma_{(p',q')}^n$ e $\gamma_{(p,q)}$ têm a mesma ordem p' e cada holonomia $\gamma_{(p',q')}^{n\alpha}$ é equivalente a $\gamma_{(p,q)}^{n'\alpha}$, eles são o mesmo conjunto de holonomias ■

Portanto, podemos sem perda de generalidade denotar os espaços-lente como $L(p, q)$. O conjunto dos espaços-lente inclui variedades de ação simples, dupla e conjunta, e constitui a maioria das formas espaciais esféricas. Embora o que se segue trate especificamente destes espaços, muitos dos resultados apresentados são extensíveis a todas as variedades esféricas. Em geral qualquer forma espacial esférica é gerada por um conjunto de geradores deste tipo (possivelmente com equadores $z_1 = 0$ e $z_2 = 0$ diferentes). Ainda segundo a classificação de [57], grupos de ação simples têm uma translação de Clifford como único gerador; grupos de ação dupla tem duas ou mais translações de Clifford como geradores; e grupos de ação conjunta têm pelo menos um gerador não-translacional que não pode ser decomposto em translações de Clifford.

A distância entre um ponto z e sua imagem por $\gamma_{(p,q)}$ é o comprimento de arco ao longo de $\mathbb{S}^3$ (parametrizada como a hipersuperfície $r = 1$ em $\mathbb{R}^4$) subtendido pela diferença entre os vetores z e $\gamma_{(p,q)}z$. A partir de (D.4d), escrevemos

$$\cos(d(z, \gamma_{(p,q)}z)) = \cos\left(\frac{2\pi}{p}\right) |z_1|^2 + \cos\left(\frac{2\pi q}{p}\right) |z_2|^2. \quad (\text{D.5})$$

Claramente, variando $|z_2|$ entre 0 e 1, temos que $r_{inj} = \pi/p \leq d(z, \gamma_{(p,q)}z)/2 \leq \pi q/p$, e que $\gamma_{(p,q)}$ é uma translação de Clifford, e portanto $L(p, q)$ é globalmente homogêneo, se e somente se $q = \pm 1$. Translações de Clifford com $q = -1$ e $q = 1$ são ditas 'translações à direita' e 'translações à esquerda', respectivamente (pois são obtidas pela multiplicação de quatérnions à esquerda e à direita). Vê-se diretamente de (D.4d) que a composição de duas translações à direita (esquerda) também é uma translação à direita (esquerda), mas a composição de uma translação à esquerda com uma à direita não é translação¹; de fato, espaços-lente gerados por ação dupla tem como geradores um par de translações de Clifford com quiralidades opostas; a ordem do grupo em tais casos é o produto das ordens dos seus subgrupos translacionais.

Para ver a 'forma' dos espaços-lente, podemos definir um fluxo substituindo $1/p \rightarrow x/p$, onde x é um parâmetro variando entre 0 e 1. As isometrias têm, então, uma forma helicoidal, com um movimento de $2\pi/p$ ao longo do eixo $z_2 = 0$ simultâneo a uma rotação (cujo sentido depende do sinal de q) de $2\pi q/p$ em torno do mesmo eixo. É possível então visualizar o domínio fundamental, que consiste da região lenticular (daí o nome) definida pela interseção entre duas grandes 2-esferas, distantes de $2\pi/p$ ao longo do eixo de simetria, e identificadas mediante uma rotação de $2\pi q/p$ em torno deste mesmo eixo. No caso de translações de Clifford, o translado e a rotação têm exatamente o mesmo valor, e a escolha do eixo é arbitrária.

O Lema D.6 garante que tudo isto vale igualmente (para possivelmente diferentes valores de p e q) para todo o grupo de isometrias gerado por $\gamma_{(p,q)}$.

A maneira como definimos $L(p, q)$ não é unívoca. Um espaço $L(p, -q)$ é a imagem no espelho de, e portanto homeomorfo à, $L(p, q)$. É imediato que $L(p, q) = L(p, rp - q)$ e portanto para restaurar a unicidade, determinamos que $0 < q < p/2$. Além disso, note que pelo teorema de Euler (D.1), se $n = q^{\phi(p)-1}$ então $\gamma_{(p,q)}^n(z_1, z_2) = (e^{2\pi i \frac{q^{\phi(p)-1}}{p}} z_1, e^{2\pi i \frac{1}{p}} z_2)$. A mera troca de z_1 por z_2 garante que os espaços $L(p, q) = L(p, q^{\phi(p)-1})$ são homeomorfos. Podemos então afirmar que

Lema D.7 *Dois espaços-lente $L(p_1, q_1)$ e $L(p_2, q_2)$ são homeomorfos se e somente se*

¹Neste caso, é sempre possível escolher (de forma única, a menos de permutação) um par de círculos máximos que sejam ao mesmo tempo equadores das duas isometrias.

$p_1 = p_2$, e também $q_1 \equiv \pm q_2 \pmod{p}$ ou $q_1 q_2 \equiv 1 \pmod{p}$.

É útil poder determinar $q^{\phi(p)-1}$ diretamente. O Lema D.4 nos permite afirmar que $q^{\phi(p)-1} = (-1)^k p_{k-1}$. Note que o Lema D.1 garante que $p_{k-1} < p/2$.

D.2.2 A geodésica fechada mais curta, e a detectabilidade da topologia

Assim como $\gamma_{(p,q)}$ gera a menor distância entre imagens topológicas para $z_2 = 0$ (ou seja, a geodésica fechada mais curta que cruza este ponto), onde $r_{inj}(z_1, 0) = r_{inj}$, $\gamma_{(p,q)}^{p_{k-1}}$ gera a menor geodésica fechada que cruza o grande círculo definido por $z_1 = 0$, igualmente com $r_{inj}(0, z_2) = r_{inj}$. Veremos que estes são casos especiais de uma regra mais geral, que permite determinar quais isometrias geram as geodésicas fechadas mais curtas em qualquer ponto da variedade, provando o seguinte teorema

Teorema D.3 *Em um espaço-lente $L(p, q)$ a menor distância entre um ponto z e sua imagem por alguma holonomia $\gamma_{(p,q)}^n z \in \Gamma_{L(p,q)}$ é gerada pela isometria $\gamma_{(p,q)}^{p_j}$, onde p_j é o denominador irredutível de algum convergente de p/q .*

Prova: A distância $d(z, \gamma_{(p,q)}^n z)$ em D.5 é minimizada se $n < p/2$ é tal que $\frac{n}{p} \pmod{1}^2$ é

mínimo ao mesmo tempo em que $\left| \pm \frac{nq}{p} \pmod{1} \right|$ é mínimo. Para qualquer $n < p$, existe j tal que $p_j < n < p_{j+1}$. Claramente $\frac{p_j}{p} \pmod{1} < \frac{n}{p} \pmod{1}$. Resta mostrar que $\frac{p_j q}{p} \pmod{1} < \frac{nq}{p} \pmod{1}$. Se $j = 1$ então por D.3 podemos escrever simplesmente

$$\frac{nq}{p} \pmod{1} = \frac{n}{(p_2 + 1/\alpha_3)}$$

. Neste caso, como $n < p_2$, claramente $\left| \pm \frac{nq}{p} \pmod{1} \right|$ é mínimo para $n = 1 = p_1$. Para $j > 1$, usamos D.3 para escrever

$$\frac{nq}{p} = n \left(\frac{q}{p} - \frac{q_j}{p_j} + \frac{q_j}{p_j} \right) = \frac{n(-1)^{j+1}}{p_j(p_{j+1} + p_j/\alpha_{j+2})} + \frac{nq_j}{p_j}$$

. Se n e j são tais que os dois termos à direita têm o mesmo sinal módulo 1, então é imediato que $n = p_j$ minimiza a soma módulo 1, e $\left| \frac{p_j q}{p} \pmod{1} \right| = \frac{1}{p_{j+1} + p_j/\alpha_{j+2}}$ é mínimo e p_j gera a geodésica mais curta. Se os termos tem sinais opostos, vamos supor por absurdo que $n \neq p_j$ gera a isometria mais curta. Note que $\left| \frac{n(-1)^{j+1}}{p_j(p_{j+1} + p_j/\alpha_{j+2})} \right| \leq \frac{1}{p^j} \leq \left| \frac{nq_j}{p_j} \right|$.

²Abusamos um pouco da notação escrevendo $\frac{r}{s} \pmod{1}$, com r/s não inteiro; mas isto denota apenas a fração menos sua parte inteira; se $r/s = [a_1, a_2, \dots, a_k]$ então $\frac{r}{s} \pmod{1} = [0, a_2, \dots, a_k]$

Assim, neste caso

$$\left| \frac{nq}{p} \bmod 1 \right| = \left| \frac{nq_j}{p_j} \bmod 1 \right| - \frac{n}{p_j(p_{j+1} + p_j/\alpha_{j+2})} = \frac{m}{p_j} - \frac{n}{p_j(p_{j+1} + p_j/\alpha_{j+2})},$$

$$\left| \frac{nq}{p} \bmod 1 \right| = \frac{p_{j+1} + [(m-1)(p_{j+1} + p_j/\alpha_{j+2}) - np_j]}{p_j(p_{j+1} + p_j/\alpha_{j+2})},$$

onde podemos tomar $m < p_j/2$, que garante que $\left| \frac{nq}{p} \bmod 1 \right| < 1/2 < 1 - \left| \frac{nq}{p} \bmod 1 \right|$. Se $m > 1$, o termo entre [parênteses] é positivo; como $p_{j+1} > p_j$, teríamos portanto $\left| \frac{nq}{p} \bmod 1 \right| > \left| \frac{p_j q}{p} \bmod 1 \right|$, que é absurdo. Portanto, a única opção restante é $m = 1$, e por consequência $nq_j \equiv 1 \bmod p_j$. Por D.4 vemos que isto implica em que $n_j \equiv p_{j-1} \bmod p_j$. Assim, escrevemos $n = bp_j + p_{j-1}$. Mas $n < p_{j+1}$, e $p_{j+1} = a_{j+1}p_j + p_{j-1}$, e portanto $r = a_{j+1} - b > 0$, e $n = p_{j+1} - rp_j$. No caso $j = k - 1$ e $r = 1$, $n = p - p_{k-1} > 1/2$, que contradiz uma das hipóteses. Nos demais casos

$$\left| \frac{nq}{p} \bmod 1 \right| = \frac{1}{p_j} - \frac{p_{j+1} - rp_j}{p_j(p_{j+1} + p_j/\alpha_{j+2})},$$

$$\left| \frac{nq}{p} \bmod 1 \right| = \frac{r + 1/\alpha_{j+2}}{(p_{j+1} + p_j/\alpha_{j+2})} > \frac{1}{p_{j+1} + p_j/\alpha_{j+2}} = \left| \frac{p_j q}{p} \bmod 1 \right|,$$

o que também é absurdo. Assim, excluímos todas as outras possibilidades, e provamos que para qualquer ponto z e inteiro $n < p$ com $p_j < n < p_{j+1}$, $\gamma_{(p,q)}^{p_j}$ gera a menor geodésica fechada. Variando n entre 1 e $p - 1$, podemos dizer portanto que em qualquer ponto z de um espaço-lente, a geodésica fechada mais curta será gerada pela holonomia $\gamma_{(p,q)}^{p_j}$, para $1 \leq j < k$ ■

Podemos assim escrever a função distância entre um ponto e sua imagem por $\gamma_{(p,q)}^{p_j}$ usando o Lema D.3

$$\cos(d(z, \gamma_{(p,q)}^{p_j} z)) = \cos\left(2\pi \frac{p_j}{p}\right) |z_1|^2 + \cos\left(\frac{2\pi}{p_{j+1} + p_j/\alpha_{j+2}}\right) |z_2|^2, \quad (\text{D.6})$$

onde usamos o fato que $\cos(x) = \cos(-x)$.

A demonstração do Teorema D.3 pode ainda ser usada para obter os geradores das sucessivas geodésicas fechadas mais curtas após a primeira. De fato,

Lema D.8 *Se em um espaço-lente $L(p, q)$ a menor distância entre um ponto z e sua imagem por alguma holonomia $\gamma_{(p,q)}^n z \in \Gamma_{L(p,q)}$ é dada pela isometria $\gamma_{(p,q)}^{p_j}$, então a segunda menor distância entre z e alguma de suas imagens topológicas é gerada por uma das seguintes isometrias: $\gamma_{(p,q)}^{p_i}$ para algum $i \neq j$, $\gamma_{(p,q)}^{2p_j}$ ou $\gamma_{(p,q)}^{p_{j+1}-p_j}$. Em particular, se $a_{j+1} = 1$ então se verifica a primeira alternativa.*

Prova: Pelo Teorema D.3, para qualquer $p_i \leq n < p_{i+1}$ para $i \neq j$, a menor geodésica fechada que passa por um ponto z liga z e sua imagem $\gamma_{(p,q)}^{p_i} z$. Para $p_j < n < p_{j+1}$,

podemos repetir exatamente os passos em D.3 para mostrar que $n = 2p_j$ gera a menor geodésica, exceto possivelmente para $n = p_{j+1} - p_j$. Se $a_{j+1} = 1$, então $p_{j+1} < 2p_j < p_{j+1}$ (e portanto $2p_j$ gera uma geodésica maior do que p_{j+1}), e $p_{j+1} - p_j = p_{j-1}$ ■

A prova acima pode ser estendida trivialmente para as geodésicas fechadas subsequentes em ordem de comprimento. Assim, se para um ponto qualquer são k o número de possíveis geradores da menor geodésica, então serão no máximo $k + c - 1$ os possíveis geradores da c -ésima menor geodésica.

É possível ainda obter uma cota independente da posição para o comprimento da menor geodésica fechada puramente em termos de p . Para valores suficientemente grandes de p , isto garante que em qualquer ponto da variedade existe pelo menos uma e no máximo duas geodésica(s) fechada(s) 'curta(s)' (i.e., pequenas em relação ao raio de curvatura) independentes.

Teorema D.4 *O raio de injetividade $r_{inj}^{L(p,q)}$ em um espaço lente $L(p, q)$, é tal que $r_{inj}^{L(p,q)} \leq \frac{\pi}{\sqrt{p}}$.*

Prova: Podemos sempre escolher $n = p_i \leq \sqrt{p} \leq p_{i+1}$. Neste caso, $\frac{2\pi p_i}{p} \leq \frac{2\pi}{\sqrt{p}}$ e $\frac{2\pi p_i q}{p} = \frac{2\pi}{p_{i+1} + p_i / \alpha_{i+2}} \leq \frac{2\pi}{\sqrt{p}}$. Para $r \leq \pi$, $\cos(r)$ é uma função decrescente de r . Já que $p_i < \frac{p}{2}$ para todo $i < k$, e $p_{i+1} > 2$, podemos ainda escrever $d(z, g^{p_i}(z)) = \cos\left(\frac{2\pi p_i}{p}\right) |Z^1|^2 + \cos\left(\frac{2\pi}{p_{i+1} + p_i / \alpha_{i+2}}\right) |Z^2|^2 \geq \cos\left(\frac{2\pi}{\sqrt{p}}\right)$. Se $p \geq 4$ então isto implica que $\frac{2\pi}{\sqrt{p}} \leq \pi$ e portanto $r_{inj}^{L(p,q)} = \frac{1}{2} \min d(z, \gamma_{(p,q)}^n z) \leq \frac{1}{2} d(z, \gamma_{(p,q)}^{p_i} z) \leq \frac{\pi}{\sqrt{p}}$. Se $p = 1, 2$ ou 3 , então $d(z, \gamma_{(p,q)}^{p_i} z) = \frac{2\pi}{p}$, e o resultado segue trivialmente ■

D.2.3 O subgrupo das holonomias translacionais

A existência de geodésicas fechadas curtas garante que em quase todos os pontos da variedade, a topologia detectável pela repetição de padrões a distâncias muito menores que o raio de curvatura será bastante parecida com translações planas. Mas é importante analisar também as isometrias que são exatamente translacionais, e determinar quando elas estão presentes em um espaço-lente, e como elas se parecem. Podemos caracterizá-las completamente:

Lema D.9 *Um espaço lente $L(p, q)$ possui translações de Clifford como holonomias se e somente se $r(q^2 + 1) \equiv 0 \pmod{p}$ para algum $r < p$. Os subgrupos translacionais terão geradores γ^n e/ou γ^m , onde $n = p/(p, q - 1)$ e/ou $m = p/(p, q + 1)$, com $n, m < p$.*

Prova: Se $\gamma_{(p,q)}^n$ é uma translação de Clifford para algum $n < p$, então $r q \equiv \pm r \pmod{p}$, e portanto $n(q - 1) \equiv 0 \pmod{p}$ e/ou $m(q + 1) \equiv 0 \pmod{p}$, onde os valores mínimos admissíveis

para n e m são $n = p/(p, q - 1)$ e $m = p/(p, q + 1)$. Podemos combinar as duas condições por multiplicação, para obter uma condição geral $r(q^2 + 1) \equiv 0 \pmod{p}$. O caso inverso é imediato. Obviamente, para $q = 1$, $n = 1$ é um gerador de translações. A condição do Lema D.6 é obedecida de imediato, com $p' = n(p', q \pm 1)$ e portanto $q = \mp 1 \pmod{p}$ ■

Note que pelo lema acima serão, no máximo, dois os subgrupos das translações de Clifford, com geradores $\gamma_{(p,q)}^n$ e $\gamma_{(p,q)}^m$, que correspondem respectivamente à multiplicação à esquerda e à direita por quatérnions $e^{2\pi i n/p}$ e $e^{2\pi i m/p}$ agindo sobre $\mathbb{S}^3$. Em espaços de ação dupla, estes são os geradores dos grupos cíclicos de quatérnions unitários que agem sobre $\mathbb{S}^3$ por multiplicação pela esquerda e direita, e correspondem à combinação dos elementos destes grupos, por um lado, e da identidade pelo outro.

Segue-se imediatamente o seguinte lema:

Lema D.10 *Um espaço-lente $L(p, q)$ não possui translações de Clifford como holonomias se p é primo e $q > 1$.*

Translações de Clifford produzem o mesmo raio de injetividade em todo o espaço, e podem portanto impor condições independentes da posição sobre a detectabilidade. O teorema abaixo mostra que existe uma conexão entre uma classe de holonomias translacionais e as holonomias mais 'curtas', e portanto mais detectáveis, geradas por γ^{p_j} .

Lema D.11 *Seja $\gamma_{(p,q)}$ um gerador de holonomias esféricas. Se $\gamma_{(p,q)}^n$ é uma translação de Clifford tal que $n \leq \sqrt{p/2}$, então $n = p_j$ para algum j .*

Prova: Pelo Lema D.9, $n = p/r$, onde $r = (p, q \pm 1)$. Pelo teorema de Lagrange, $n = p_j$ se a quantidade $Q = 2n|q - ar| - r$ for negativa para algum $a \in \mathbb{N}$. Mas como $(q, r) = 1$, $|q - ar| \geq 1$. Tomando $a = q \pm 1/r$ (que por construção é inteiro), temos $|q - ar| = 1$, e portanto mínimo. Assim, com esta escolha $Q = 2n - r = 2n - p/n$; o que implica que Q é negativo, e portanto $\gamma_{(p,q)}^n$ é uma translação de Clifford, se $n \leq \sqrt{p/2}$ ■

Bibliografia

- [1] Abramo, L.R., Bernui, A., Ferreira, I., Villela, T. & Wuensche, C.A., *Phys. Rev.* **D74**, 063506 (2006).
- [2] Agol, I., Dunfield, N.M., Storm, P.A. & Thurston, W., *Lower bounds on volumes of hyperbolic Haken 3-manifolds* (2005), arXiv:math.DG/0506338.
- [3] Alcaniz, J.S. & Pires, N., *Phys. Rev.* **D70**, 047303 (2004).
- [4] Allen, S.W. et al., *Mon. Not. R. Astron. Soc.* **353**, 457 (2004).
- [5] Alpher, R.A., Bethe, H. & Gamow, G., *Physical Review* **73**, 803 (1948). Hans Bethe não teve absolutamente nenhuma participação neste trabalho. O seu nome foi adicionado por razões puramente eufônicas por Gamow.
- [6] Aurich, R., Lustig, S. & Steiner, F., *Class. Quantum Grav.* **22**, 2061 (2005).
- [7] Aurich, R., Lustig, S. & Steiner, F., *Class. Quantum Grav.* **22**, 3443 (2005).
- [8] Aurich, R., Lustig, S. & Steiner, F., *Mon. Not. Roy. Astron. Soc.* **369** 240 (2006).
- [9] Barris, B.J. et al., *Astrophysical Journal* **602**, 571 (2004).
- [10] Bennett, C.L., Bay, M., Halpern, M. et al., *Astrophysical Journal* **583**, 1,(2003).
- [11] Bennett, C.L., Halpern, M., Hinshaw, G. et al., *Astrophysical Journal supp.* **148**, 1 (2003).
- [12] Bento, M.C., Bertolami, O. & Sen, A.A., *Phys. Rev.* **D66**, 043507 (2002).
- [13] Bento, M.C., Bertolami, O., Rebouças, M.J. & Santos, N.M.C., *Phys. Rev.* **D73**, 103521 (2006).
- [14] Bernui, A. & Villela, T., *Astron. Astrophys.* **445**, 795 (2006).
- [15] Bernui, A., *Brazilian Journal of Physics* **35**, 1185 (2005).

- [16] Bernui, A., Mota, B., Rebouças, M.J. & Tavakol, R. *Astron. Astrophys.* **464** (2007).
- [17] Bernui, A., Mota, B., Rebouças, M.J. & Gomero, G.I., *Proceedings of the 10th Marcel Grossmann Meeting*, Recent developments in theoretical and experimental general relativity, gravitation, and relativistic field theories **C**, 1657, Rio de Janeiro (2003).
- [18] Bernui, A., Tsallis, C. & Villela, T., *Phys. Lett.* **A356**, 426 (2006).
- [19] Bernui, A., Villela, T., Wuensche, C. A., Leonardi, R. & Ferreira, I., *Astron. Astrophys.* **454**, 409 (2006).
- [20] Bertolami, O., Sen, A.A., Sen, I. & Silva, P.T., *Mon. Not. Roy. Astron. Soc.* **353**, 329 (2004).
- [21] Bertone, G., Hooper, D. & Silk, J., *Phys. Rep.* **405**, 279 (2005).
- [22] Bielewicz, P., Eriksen, H. K., Banday, A. J., Gorski, K. M. & Lilje, P. B., *Astrophysical Journal* **635**, 750 (2005).
- [23] Bilic, N., Tupper, G.B. & Viollier, R.D., *Phys. Lett.* **B535**, 17 (2002).
- [24] Bond, J.R., Pogosyan, D. & Souradeep, T., *Phys. Rev.* **D62**, 043005 (2000).
- [25] Boothby, W.M., *An introduction to differentiable manifolds and Riemannian geometry*, 2nd Edition, London, Academic Press (1986).
- [26] Bunn, E.F. & Scott, D., *Mon. Not. Roy. Astron. Soc.* **313**, 331 (2000).
- [27] Burbidge, E.M., Burbidge, G.R., Fowler, W.A. & Hoyle, F., *Reviews of Modern Physics* **29**, 547 (1957).
- [28] Burles, S., Nollett, K.M. & Turner, M.S., *Astrophysical Journal* **552**, L1 (2001).
- [29] Calvão, M.O., Gomero, G.I., Mota, B. & Rebouças, M.J., *Class. Quantum Grav.* **22**, 1991 (2005).
- [30] Chiang, L.-Y., Naselsky, P.D., Verkhodanov, O.V. & Way, M.J., *Astrophysical Journal* **590**, L65 (2003).
- [31] Choudhury, T.R. & Padmanabhan, T., *Mon. Not. Roy. Astron. Soc.* **344**, 823 (2003).
- [32] Coles, P., Dineen, P., Earl, J. & Wright, D., *Mon. Not. Roy. Astron. Soc.* **350**, 983 (2004).

- [33] Cooray, A. & Seto, N. (2005), astro-ph/0509039.
- [34] Copi, C.J., Huterer, D. & Starkman, G.D. *Phys. Rev.* **D70**, 043515 (2004).
- [35] Copi, C.J., Huterer, D., Schwarz, D.J. & Starkman, G.D., *Mon. Not. Roy. Astron. Soc.* **367**, 79 (2006).
- [36] Copi, C.J., Huterer, D., Schwarz, D.J. & Starkman, G.D., *Phys. Rev.* **D75**, 023507 (2007).
- [37] Cornish, N.J., *Phys. Rev.* **D63**, 027302 (2000).
- [38] Cornish, N.J., Spergel, D. N., Starkman, G.D. & Komatsu, E., *Phys. Rev. Lett.* **92**, 201302 (2004).
- [39] Cornish, N.J., Spergel, D.N. & Starkman, G.D., *Class. Quantum Grav.* **15**, 2657 (1998).
- [40] Coxeter, H.S.M., *Non-Euclidean Geometry* 5th ed., Toronto, University of Toronto Press (1965).
- [41] Crotty, P., Garcia-Bellido, J. , Lesgourgues, J. & Riazuelo, A., *Phys. Rev. Lett.* **91**, 171301 (2003).
- [42] da Costa, L.N., et al., *Mon. Not. Roy. Astron. Soc.* **299**, 425 (1998).
- [43] de Oliveira-Costa, A. & Tegmark M., *Phys. Rev.* **D74**, 023005 (2006).
- [44] de Oliveira-Costa, A., Tegmark, M., Zaldarriaga, M. & Hamilton, A., *Phys. Rev.* **D69**, 063516 (2004).
- [45] Dev, A., Jain, D. & Alcaniz, J. S., *Astron. Astrophys.* **417**, 847 (2004).
- [46] Dicke, R.H., Peebles, P.J.E., Roll, P.G. & Wilkinson, D.T., *Astrophysical Journal* **142**, 414 (1965).
- [47] d’Inverno, R., *Introducing Einstein’s Relativity*, Oxford Univerity Press, Oxford, (1992).
- [48] Donoghue, E.P. & Donoghue, J. F., *Phys. Rev.*, **D71**, 043002 (2005).
- [49] Ellis, G.F.R. & Van Elst, H., *Cosmological Models*, Cargèse Lectures (1998), gr-qc/9812046.
- [50] Ellis, G.F.R., Topology and cosmology. *Gen. Rel. Grav.* **2**, 1 (1971).

- [51] Eriksen, H.K., Banday, A. J., Górski, K.M. & Lilje, P.B., *Astrophysical Journal* **612**, 633 (2004).
- [52] Eriksen, H.K., Banday, A. J., Górski, K.M. & Lilje, P.B., *Astrophysical Journal* **622**, 58 (2005).
- [53] Eriksen, H.K., Hansen, F. K., Banday, A.J., Gorski, K.M. & Lilje, P. B., *Astrophysical Journal* **605**, 14 (2004).
- [54] Feldman, H. A. et al., *Astrophysical Journal* **596**, L131 (2003).
- [55] Freedman, W., *Astrophysical Journal* **553**, 47 (2001).
- [56] Gabai, D., Meyerhoff, R. & Milley, P., *Minimum volume cusped hyperbolic three-manifolds* (2007), arXiv:0705.4325.
- [57] Gausmann, E. ,Lehoucq, R., Luminet, J.-P., Uzan, J.-P. & Weeks, J., *Class. Quantum Grav.* **18**, 5155 (2001).
- [58] Gomero, G.I., Rebouças, M.J. & Teixeira, A.F.F., *Phys. Lett.* **A275**, 355 (2000).
- [59] Gomero, G.I., Rebouças, M.J. & Teixeira, A.F.F., *Class. Quantum Grav.* **18**, 1885 (2001).
- [60] Gomero, G.I., Teixeira, A. F. F., Rebouças, M.J. & Bernui, A., *Int. J. Mod. Phys.* **D11**, 869 (2002).
- [61] Gomero, G.I. & Rebouças, M.J., *Phys. Lett.* **A311**, 319 (2003).
- [62] Gomero, G.I. *Busca da topologia do Universo usando fontes cósmicas discretas*, CBPF, Tese de Doutorado (2002).
- [63] Gomero, G.I., Rebouças, M.J. & Tavakol, R., *Class. Quantum Grav.* **18**, 4461 (2001).
- [64] Gomero, G.I., Rebouças, M.J. & Tavakol, R., *Class. Quantum Grav.* **18**, L145 (2001).
- [65] Gomero, G.I., Rebouças, M.J. & Tavakol, R., *Int. J. Mod. Phys.* **A17**, 4261 (2002).
- [66] Gordon, C., Hu, W., Huterer, D. & Crawford, T. *Phys. Rev.* **D72**, 103002 (2005).
- [67] Gorski, K.M., Hivon, E., Banday, A.J. et al. *Astrophysical Journal* **622**, 759 (2005).

- [68] Gundermann, J. (2005), *Predicting the CMB power spectrum for binary polyhedral spaces*, astro-ph/0503014
- [69] Guth, A.H., *Phys. Rev.* **D23**, 347 (1981).
- [70] Hajian, A. & Souradeep, T. (2005), astro-ph/0501001.
- [71] Hajian, A. & Souradeep, T., *Astrophysical Journal* **597**, L5 (2003).
- [72] Hajian, A., Souradeep, T. & Cornish, N., *Astrophysical Journal* **618**, L63 (2004).
- [73] Hansen, F.K., Banday, A. J. & Górski, K.M., *Mon. Not. Roy. Astron. Soc.* **354**, 641 (2004).
- [74] Hansen, F.K., Cabella, P., Marinucci, D. & Vittorio, N., *Astrophysical Journal* **607**, L67 (2004).
- [75] Hatcher, A., *Algebraic Topology*, Cambridge, Cambridge University Press (2001).
- [76] Hillebrandt, W. & Niemeyer, J.C., *Annual Review of Astronomy and Astrophysics* **38**, 191 (2000).
- [77] Hinshaw, G., Nolta, M., R., Bennett, C.L. et al., (2006), astro-ph/0603451.
- [78] Hinshaw, G., Spergel, D.N., Verde, L. et al., *Astrophysical Journal supp.* **148**, 135 (2003).
- [79] Hipolito-Ricaldi, W.S. & Gomero, G.I., *Phys. Rev.* **D72**, 103008 (2005).
- [80] Hodgson, C.D. & Weeks, J., *Experimental Mathematics* **3**, 261 (1994).
- [81] Hu, W., *CMB Anisotropies: a decadal survey*, "Birth and Evolution of the Universe", (1999), astro-ph/0002520.
- [82] Levin, J., *Phys. Rev.* **D70**, 083001 (2004).
- [83] Kamenshchik, A., Moschella, U. & Pasquier, V., *Phys. Lett.* **B511**, 265 (2001).
- [84] Khinchin, A.Ya., *Continued fractions*, University of Chicago Press, Chicago, EUA (1964).
- [85] Kirkman, D., Tytler, D., Suzuki, N.J., O'Meara, M. & Lubin, D., *Astrophysical Journal supp.* **149**, 1 (2003).
- [86] Knop, R.A. et al., *Astrophysical Journal* **598**, 102 (2003).

- [87] Komatsu, E., Kogut, A., Nolte, M. et al., *Astrophysical Journal supp.* **148**, 119 (2003).
- [88] Lachièze-Rey, M. & Luminet, J.-P., *Phys. Rep.* **254**, 135 (1995).
- [89] Land, K. & Magueijo, J., *Mon. Not. Roy. Astron. Soc.* **357**, 994 (2005).
- [90] Land, K. & Magueijo, J., *Phys. Rev. Lett.* **95**, 071301 (2005).
- [91] Larson, D.L. & Wandelt, B.D., *Astrophysical Journal* **613**, L85 (2004).
- [92] Lehoucq, R., Lachièze-Rey, M. & Luminet, J.-P., *Astron. Astrophys.* **313**, 339 (1996).
- [93] Levin, J., *Phys. Rep.* **365**, 251 (2002).
- [94] Liddle, A.R., *An introduction to cosmological inflation*, (1999) astro-ph/9901124.
- [95] Lima, E.L., *Grupo fundamental e espaços de recobrimento*, Projeto Euclides, Rio de Janeiro, IMPA (2006).
- [96] Linde, A., *Phys. Lett.* **B108**, 389 (1982).
- [97] Linde, A., *Prospects for inflation*, Extended version of the talk at the Nobel Symposium, (2003), hep-th/0402051.
- [98] Luminet, J.-P., Weeks, J., Riazuelo, A., Lehoucq, R. & Uzan, J.-P., *Nature* **425**, 593 (2003).
- [99] Makler, M., *Gravitational Dynamics of Structure Formation in the Universe*, Tese de Doutorado, CBPF (2001).
- [100] Makler, M., Mota, B. & Rebouças, M.J., *Proceedings of the 10th Marcel Grossmann Meeting*, General Relativity **B**, 1117, Rio de Janeiro (2003).
- [101] Makler, M., Mota, B. & Rebouças, M.J., *Constraints from the detection of cosmic topology on the generalized Chaplygin gas* (a ser submetido a publicação) (2005), astro-ph/0507116.
- [102] Makler, M., Oliveira, S.Q. & Waga, I., *Phys. Lett.* **B555**, 1 (2003).
- [103] Makler, M., Oliveira, S.Q. & Waga, I., *Phys. Rev.* **D64**, 123521 (2003).
- [104] Marion, J.B., *Classical Dynamics of Particles and Systems*, Academic Press Inc., New York (1970).

- [105] Massey, W. S., *Algebraic Topology: An Introduction*, Graduate Texts in Mathematics 56, New York, Springer-Verlag (1977).
- [106] Mathers, J.C. et al., *Astrophysical Journal* **512**, 511 (1999).
- [107] McCabe, G., *The Structure and Interpretation of Cosmology* (2003),
<http://philsci-archive.pitt.edu/archive/00001507/>.
- [108] McInnes, B., *Nucl. Phys.* **B709**, 213 (2005).
- [109] Menzies, D. & Mathews, G.J., *Astrophysical Journal* **624**, 7 (2005).
- [110] Miller, D., *The Integrated Sachs-Wolfe Effect: Origin, Status and Future*, (1999),
http://cfcp.uchicago.edu/davemilr/ISW/ISW_effect_Home.
- [111] Misner, C.W., Wheeler, J.A. & Thorne, K., *General Relativity*, University of Chicago Press, Chicago (1984).
- [112] Mostow, G.D. *Ann. Math. Studies* **78**, Princeton University Press, Princeton, EUA (1973).
- [113] Mota, B., Bernui, A., Gomero, G. I. & Rebouças, M. J., *Injectivity profiles for flat 3-manifolds* (não publicado).
- [114] Mota, B., Bernui, A., Rebouças, M.J. & Tavakol, R., *Int. J. Mod. Phys.* **D16**, 411 (2007).
- [115] Mota, B., Gomero, G. I., Rebouças, M.J. & Tavakol, R., *Class. Quantum Grav.* **21**, 3361 (2004).
- [116] Mota, B., Makler, M. & Rebouças, M.J., *Astron. Astrophys.* **446**, 805 (2006).
- [117] Mota, B., Rebouças, M.J. & Tavakol, R., *Class. Quantum Grav.* **20**, 4837 (2003).
- [118] Mota, B., Rebouças, M.J. & Tavakol, R., *Int. J. Mod. Phys.* **A20**, 2415 (2005).
- [119] Nesseris, S. & Perivolaropoulos, L., *Phys. Rev.* **D70**, 043531 (2004).
- [120] Nieto, M.M. & Turyshv, S.G., *Class. Quantum Grav.* **21**, 4005 (2004).
- [121] Oliveira Santos, J.-P., *Introdução à Teoria dos Números*, Coleção Matemática Universitária, IMPA, Rio de Janeiro (1998).
- [122] Olofsson, P., *Probability, Statistics, and Stochastic Processes*, Wiley-Interscience, Hoboken, New Jersey (2005).

- [123] O’Mullane, W., Banday, A. J., Gorski, K. M., Kunszt, P., Szalay, A. S. , Proceedings of the MPA/ESO/MPE Workshop. ISBN 3-540-42468-7. Edited by A.J. Banday, S. Zaroubi, and M. Bartelmann. Springer-Verlag, (2001).
- [124] Ozols, V., *Proceedings of the American Mathematical Society* **44**, 1 (1974).
- [125] Park, C.-G., *Mon. Not. Roy. Astron. Soc.* **349**, 313 (2004).
- [126] Park, C.-G., Park, C. & Gott III, J. R., *Astrophys. Journal* **660**, 959 (2007).
- [127] Partridge, R.B., *Rencontres de Moriond*, (2002),
<http://moriond.in2p3.fr/J02/Talks2002/B.Partridge/LesArcs.ps>.
- [128] Peebles, P.J.E., *Principles of physical cosmology*, Princeton, Princeton University Press (1993).
- [129] Peebles, P.J.E. & Ratra, B. *Rev. Mod. Phys.* **75**, 559 (2003).
- [130] Penrose, R., *Proc. Cambridge Philos. Soc.* **55**, 137 (1959).
- [131] Penzias, A. A. & Wilson, R. W., *Astrophysical Journal* **142**, 419 (1965).
- [132] Perlmutter, S. & Schmidt, B. P., *Lecture Notes in Physics* **598**, 195 (2003).
- [133] Perlmutter, S. et al., *Nature* **391**, 51 (1998).
- [134] Poincaré, H., *Acta Mathematica* **III**, 49 (1883).
- [135] Pope, A.C. et al., *Astrophysical Journal* **607**, 655 (2004).
- [136] Prunet, S., Uzan, J.-P., Bernardeau, F. & Brunier, T., *Phys. Rev.* **D71**, 083508 (2005).
- [137] Kar, S., SenGupta, S., *The Raychaudhuri equations: a brief review*, a ser publicado no Pramana, (2006), astro-ph/0611123.
- [138] Rebouças, M.J., *A brief introduction to cosmic topology*, Proceedings of the XIth school of Cosmology and Gravitation, editado por M. Novello e S.E. Perez Bergliaffa, 188 (2005).
- [139] Rebouças, M.J., Alcaniz, J.S., Mota, B. & Makler, M., *Astron. Astrophys.* **452**, 803 (2006).
- [140] Reed, M. & Simon, B., *Functional Analysis I*, Methods of Modern Mathematical Physics, San Diego, Academic Press (1980).

- [141] Reis, R.R.R., Waga, I., Calvão, M.O. & Jorás, S.E., *Phys. Rev.* **D68**, 061302(R) (2003). [astro-ph/0306004](#)
- [142] Riazuelo, A., Uzan, J.-P., Lehoucq, R. & Weeks, J., *Phys. Rev.* **D69**, 103514 (2004).
- [143] Riazuelo, A., Weeks, J., Uzan, J.-P., Lehoucq, R. & Luminet, J.-P., *Phys. Rev.* **D69**, 103518 (2004).
- [144] Riess, A.G. et al., *Astrophys. Journal* **607**, 665 (2004).
- [145] Riess, A.G. et al., *Astrophysical Journal* **116**, 1009 (1998).
- [146] Rindler, W., *Essential Relativity*, revised second edition, Chapter 3, Springer-Verlag, Heidelberg (1979).
- [147] Roukema B.F. & Edge A., *Mon. Not. Roy. Astron. Soc.* **292**, 105 (1997).
- [148] Roukema, B.F., Lew, B., Cechowska, M., Marecki, A. & Bajtlik, S., *Astron. Astrophys.* **423**, 821 (2004).
- [149] Schwarz, D.J., Starkman, G.D., Huterer, D. & Copi, C.J., *Phys. Rev. Lett.* **93**, 221301 (2004).
- [150] Scott, D. & Smoot, G.F., *J. Phys.* **G33**, 1 (2006).
- [151] Shapiro Key, J., Cornish, N.J., Spergel, D.N. & Starkman, G.D., *Phys. Rev.* **D75**, 084034 (2007).
- [152] Silva e Costa, S., *Phys. Lett* **A328**, 270 (2004).
- [153] Smoot, G.F., *Information for the public*, Royal Swedish Academy of Sciences, (2006), http://nobelprize.org/nobel_prizes/physics/laureates/2006/info.html.
- [154] Souradeep, T. & Hajian, A. (2005), [astro-ph/0502248](#).
- [155] Souradeep, T. & Hajian, A., *Pramana* **62**, 793 (2004).
- [156] Spergel, D.N., Bean, R., Doré, O. et al. (2006), [astro-ph/0603449](#).
- [157] Spergel, D.N., Verde, L., Peiris, H.V. et al., *Astrophysical Journal supp.* **148**, 175 (2003).
- [158] Spergel, D.N., et al., Wilkinson Microwave Anisotropy Probe (WMAP). Three Year Results: Implications for Cosmology (2006), [astro-ph/0603449](#) (a ser publicado no *Astrophys. J.*).

- [159] Spergel, D.N., *et al.*, *Astrophysical Journal* **148**, 175 (2003).
- [160] Starkman, G.D., *Class. Quantum Grav.*, **15**, 2529 (1998).
- [161] Starobinski, A.A., *Phys. Lett.* **B91**, 99 (1980).
- [162] Stolyarov, V. et al., *Mon. Not. Roy. Astron. Soc.* **336**, 97 (2002),
<http://www.rssd.esa.int>.
- [163] Tegmark, M., *J. Cosmol. Astropart. Phys.* **04**, 01 (2005)
- [164] Tegmark, M., de Oliveira-Costa, A. & Hamilton, A. J. S., *Phys. Rev.* **D68**, 123523 (2003).
- [165] Tegmark, M. et al., *Phys. Rev.* **D69**, 103501 (2004).
- [166] Teixeira, A.F.F. (2003), physics/0312013.
- [167] Threlfall, W. & Seifert, H. *Math. Annalen* **104**, 543 (1932).
- [168] Thurston, W.P. & Weeks, J., *Scientific American* **July**, 94 (1984).
- [169] Thurston, W.P., *Three-Dimensional Geometry and Topology. Vol.1*, Edited by Silvio Levy, Princeton University Press (1997).
- [170] Tolman, R.C., *Relativity, Thermodynamics and Cosmology*, Dover Publications, New York, (1987).
- [171] Tomita, K. (2005), astro-ph/0509518.
- [172] Tonry, J.L. et al., *Astrophysical Journal* **594**, 1 (2003).
- [173] Uzan, J.-P., Lehoucq, R. & Luminet, J.-P., *Astron. Astrophys.* **351**, 776 (1999).
- [174] Vale, C. (2005), astro-ph/0510137.
- [175] Vielva, P., Martínez-González, E., Barreiro, R. B., Sanz, J. L. & Cayon, L. , *Astrophysical Journal* **609**, 22 (2006).
- [176] Weeks, J., *SnapPea: A computer program for creating and studying hyperbolic 3-manifolds*, <http://thames.northnet.org/weeks/>
- [177] Weeks, J., Lehoucq, R. & Uzan, J.-P., *Class. Quantum Grav.* **20**, 1529 (2003).
- [178] Weeks, J., *Curved Spaces*, <http://www.geometrygames.org/CurvedSpaces/index.html>

- [179] Weeks, J., *Detecting topology in a nearly flat hyperbolic universe*, (2002), [astro-ph/0212006](#).
- [180] Weeks, J., *Hyperbolic structures on 3-manifolds*, Princeton Univ. Ph.D. thesis (1985).
- [181] Weeks, J., *Maxwell's Multipole Vectors and the CMB*, (2004), [astro-ph/0412231](#).
- [182] Weeks, J., *Mod. Phys. Lett. A* **18**, 2099 (2003).
- [183] Weeks, J., Lehoucq, R. & Uzan, J.-P., *Class. Quantum Grav.* **20**, 1529 (2003).
- [184] White, M. & Hu, W., *Astronomy and Astrophysics* **321**, 8 (1997).
- [185] Wiaux, Y., Vielva, P., Martínez-González, E. & Vandergheynst, P., *Phys. Rev. Lett.* **96**, 151303 (2006).
- [186] Wibig, T. & Wolfendale, A. W., *Mon. Not. Roy. Astron. Soc.* **360**, 236 (2001).
- [187] Willick, J. & Strauss, M., *Astrophysical Journal* **507**, 64 (1998).
- [188] WMAP homepage, <http://map.gsfc.nasa.gov/>
- [189] Wolf, J.A., *Spaces of constant curvature*, London, McGraw-Hill (1967).
- [190] Zeldovich, Ya. B. & Novikov, I.D., *The Structure and Evolution of the Universe*, The University of Chicago Press (1983).
- [191] Zhu, Z.-H., *Astron. Astrophys.* **423**, 421 (2004).