TROISIÈME CYCLE DE LA PHYSIQUE

EN SUISSE ROMANDE

New methods for the determination of cosmological parameters

by

Ruth Durrer[1] and Norbert Straumann[2]

[1]Département de Physique Théorique, Université de Genève
Quai E. Ansermet 24, 1211 Genève 4, Switzerland
[2] Institut für Theoretische Physik Universität Zurich
Winterthurerstrasse 180, 8057 Zürich, Switzerland

1999

TROISIÈME CYCLE DE LA PHYSIQUE
EN SUISSE ROMANDE

New methods for the determination of cosmological parameters

by

Ruth Durrer[1] and Norbert Straumann[2]

[1]Département de Physique Théorique, Université de Genève
Quai E. Ansermet 24, 1211 Genève 4, Switzerland
[2] Institut für Theoretische Physik Universität Zurich
Winterthurerstrasse 180, 8057 Zürich, Switzerland

1999

TROISIÈME CYCLE DE LA PHYSIQUE

EN SUISSE ROMANDE

UNIVERSITÉS DE :

BERNE - FRIBOURG - GENÈVE

LAUSANNE - NEUCHÂTEL

ET

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

ARCHIVES ET VENTE DE POLYCOPIÉS

M. D. Reymond, Université de Lausanne
Bâtiment des Sciences physiques
1015 LAUSANNE-DORIGNY

Nouvelles méthodes de détermination des paramètres cosmologiques

Durant ces vingt dernières années, la cosmologie est passé du statut de science très qualitative et spéculative à celui de véritable science quantitative. Récemment cette évolution s'est considérablement accentué par l'important développement des nouvelles méthodes de détermination du paramètre de décélération (ou plus précisément d'accélération) à partir des supernovae de type Ia, des études de l'effet Sunyaev-Zel'dovich dans les amas de galaxies, de l'analyse des lentilles gravitationnelles ou des observations du fond cosmique micro-onde.

Dans ce cours, nous étudions dans la première partie les lentilles gravitationnelles et dans la seconde partie les anisotropies du fond cosmique micro-onde. Nous avons souhaité montrer d'une part que ces sujets permettent un très attrayant traitement théorique, et d'autre part souligner que, sous certaines hypothèses, les données expérimentales prévues les prochaines années permettrons de déterminer à quelques pourcents près les paramètres cosmologiques tels que la courbure de l'espace, le paramètre de Hubble ou la constante cosmologique, une nouveauté inouie en cosmologie.

Nous tenons à remercier notre assistant Martin Kunz pour son aide très efficace concernant la préparation des notes de ce cours, et pour sa lecture soigneuse de la première version de ce texte. Nos remerciements vont aussi à Lukas Grenacher. Enfin nous remercions le Troisième Cycle de la Suisse Romande de nous avoir invité à présenter les nouveaux développements de ce fascinant sujet.

Ruth Durrer et Norbert Straumann

Contents

Part I

Lectures on Gravitational Lensing

Introduction

Gravitational lensing has become one of the important fields in present day astronomy. The enormous activity in this area has largely been driven by considerable improvements of the observational capabilities. So far we have, however, only scratched the surface. The rate and quality of the data will increase dramatically, thanks to new wide–field cameras and imaging with new telescopes, in particular those of the 8m–class.

Why is gravitational lensing so important? The answer is simple: It has the distinguishing feature of being independent of the nature and the physical state of the deflecting mass. Therefore, it is perfectly suited to study dark matter at all scales.

Let me mention, for illustration, just one topic which has recently attracted a lot of attention. This concerns the parameter–free reconstructions of projected mass distributions from (weak) lensing data, for instance, for rich clusters of galaxies. Fig. 1 shows an example of such data, which has been obtained with the wide–field camera (WFPC 2) on HST. Beside arcs one can see many arclets which are weakly distorted images of faint distant galaxies. We shall see in these lectures how one analyses such data.

It is always interesting to know something about the history of a field. I will inject historical remarks at appropriate places here and there during the course. Let me now mention, however, that Einstein discovered gravitational lensing as early as in 1912, before the general theory of relativity was formulated. At the time he was working on the static limit of a relativistic theory of gravity. Reconstruction of some of Einstein's research notes dating back to 1912 reveals that he explored the possibility of gravitational lensing. These research notes can be found in [1], see also [2].) Einstein did the gravitational lensing calculations during a visit to Berlin where he met the astronomer Freundlich at the Königliche Sternwarte (Royal Observatory). Beside considering the possibility of a double image of a source as a result of gravitational bending, he also computed the magnification of the intensity of these images. As is well known, it is this effect which the present day MACHO search relies upon.

The first who recognized the great potential of gravitational lensing was Fritz Zwicky back in 1937. In two short and very impressive papers [3] he pointed out that galaxies can split images of background sources by a large enough angle to be observed. I shall later go through Zwicky's papers. As you will see, virtually all of his predictions have come true (about 50 years later).

Let me add a few remarks on the program of these lectures. I want to derive already at the beginning - as directly as possible - all the important general lensing equations. For this I need one simple consequence of general relativity (GR), namely that one can reduce gravitational lensing to a problem in ordinary ray optics, with an effective refraction index that is simply given in terms of the Newtonian potential (for an almost Newtonian situation, *i.e.*, weak fields). Those of you who do not know GR should accept this fact.

The table of contents gives a more detailed description of the program.

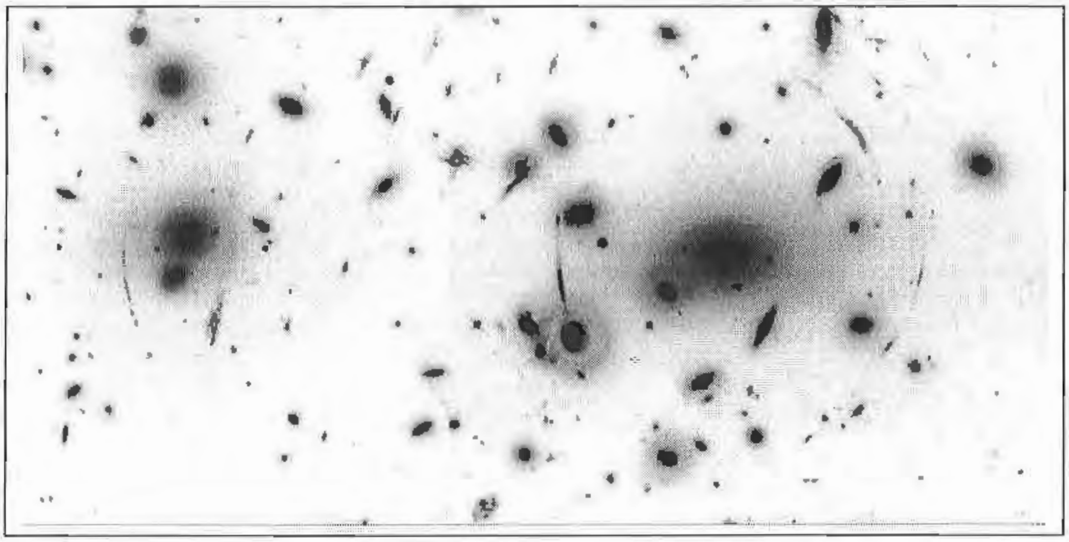

Figure 1: Hubble Space Telescope image of the cluster Abell 2218. Beside arcs around the two centers of the cluster, many arclets can be seen (NASA HST Archive).

Chapter 1

Basic lensing equations

The conceptual basis of gravitational lensing is extremely simple. At the same time this is the main reason why it is so important for the study of mass distributions at all scales. For all practical purposes we can use the ray approximation for the description of light propagation. In this limit, the rays correspond to null geodesics in a given gravitational field, and the polarization vector obeys the law of parallel transport.

These laws can be deduced from Maxwell's equations; see, e.g., [4], Section 1.8. Let us briefly recapitulate the eikonal approximation for the Maxwell field F. As usual we set

$$F = fe^{iS} \tag{1.1}$$

with a slowly varying f and a real S. (I omit to indicate that the real part on the right has to be taken.) From $dF = 0$ we get

$$df + if \wedge dS = 0 \tag{1.2}$$

and $d * F = 0$ implies

$$d * f + i(*f) \wedge dS = 0. \tag{1.3}$$

In these equations we neglect the differentials of f and $*f$,

$$f \wedge dS = 0, \; (*f) \wedge dS = 0. \tag{1.4}$$

These relations imply

$$0 = i_{\nabla S}(f \wedge dS) = (i_{\nabla S} f) \wedge dS + f\,(\nabla S)^2\,. \tag{1.5}$$

Since the second equation in (1.4) is equivalent to $i_{\nabla S} f = 0$, we obtain the general relativistic eikonal equation

$$(\nabla S)^2 = 0 \quad (g^{\mu\nu}\partial_\mu S\,\partial_\nu S = 0)\,. \tag{1.6}$$

Let me also repeat the Hamilton-Jacobi method for light propagation. Consider rays $\gamma(\lambda)$, i.e., trajectories orthogonal to the wave fronts $\{S = \text{const.}\}$:

$$\dot{\gamma}(\lambda) = \nabla S\,(\gamma\,(\lambda))\,. \tag{1.7}$$

We show that $\gamma(\lambda)$ is a null geodesic, λ an affine parameter. First, $\dot{\gamma}$ is a null vector,

$$g\,(\dot{\gamma}, \dot{\gamma}) = g\,(\nabla S, \nabla S) = 0. \tag{1.8}$$

Secondly, we have

$$\nabla_{\dot{\gamma}}\dot{\gamma} = \nabla_{\nabla S}\,(\nabla S)\,\big|_{\gamma(\lambda)}\,. \tag{1.9}$$

Here, the right hand side is

$$(\nabla^\nu S)\,\nabla_\nu\,(\nabla^\mu S) = \nabla^\nu S\,\nabla^\mu\,(\nabla_\nu S) = \frac{1}{2}\nabla^\mu\,(\nabla^\nu S\,\nabla_\nu S) = 0. \tag{1.10}$$

For sufficiently strong lenses the wave fronts develop edges and self-intersections (see Fig. 1.1). An observer behind such folded fronts obviously sees more than one image. From Fig. 1.1 one also understands how the time delay of pairs of images arises: this is just the time elapsed between crossings of different sheets of the same wave front. Hopefully, this will lead, for instance, to an accurate determination of the Hubble constant (see Section 4.2).

Fig. 1.1 also shows, where strong and weak lensing occurs behind a cluster of galaxies.

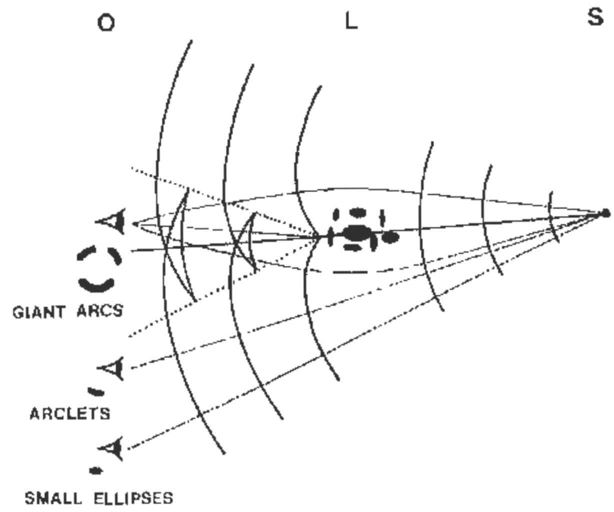

Figure 1.1: Wave fronts in the presence of a cluster perturbation.

1.1 Reduction to a problem of ordinary ray optics, effective refraction index

For the time being, we consider almost Newtonian, asymptotically flat situations. Generalizations to the cosmological context are easy and basically amount to interpreting all distances in the formulas derived below as angular diameter distances. GR implies then that gravitational lensing theory is just usual ray optics with the effective refraction index

$$n(\mathbf{x}) = 1 - 2U(\mathbf{x})/c^2, \tag{1.11}$$

where $U(\mathbf{x})$ is the Newtonian potential of the mass distribution $\rho(\mathbf{x})$,

$$U(\mathbf{x}) = -G \int \frac{\rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x'. \tag{1.12}$$

Until we come to cosmological problems, this is the only fact which you have to accept from GR. For those of you who have some knowledge of this great theory, I describe briefly how one arrives at (1.11).

For an almost Newtonian situation, the metric element of spacetime (in almost Lorentzian coordinates) is given by

$$g = \left(1 + 2U/c^2\right) dt^2 - \left(1 - 2U/c^2\right) d\mathbf{x}^2 \tag{1.13}$$

(see, e.g., [4], Section 4.2).

On the other hand, the spatial part of a light ray satisfies Fermat's principle

$$\delta \int \frac{d\sigma}{\sqrt{g_{00}}} = 0, \tag{1.14}$$

for variations with fixed end points ([4], Section 1.7). Here $d\sigma^2$ denotes the spatial part of (1.13). Hence, we arrive at Fermat's principle in ordinary ray optics

$$\delta \int n\left(\mathbf{x}\left(s\right)\right)\left|\dot{\mathbf{x}}\left(s\right)\right| ds = 0, \tag{1.15}$$

where the refraction index n is given by (1.11); $\mathbf{x}(s)$ is the parametrized light path and $\left|\dot{\mathbf{x}}\left(s\right)\right|$ denotes the euclidean norm of the tangent vector.

Let me remind you how one obtains from Fermat's principle (1.15) the basic ray equation in optics. You can regard (1.15) as a Hamiltonian variational principle, with Lagrangian

$$L\left(\mathbf{x}, \dot{\mathbf{x}}\right) = n(\mathbf{x})\sqrt{\dot{\mathbf{x}}^2}. \tag{1.16}$$

When s is the euclidean path length parameter, we have:

$$\frac{\partial L}{\partial \mathbf{x}} = \boldsymbol{\nabla} n, \qquad \left(\dot{\mathbf{x}}^2 = 1\right);$$

$$\frac{\partial L}{\partial \dot{\mathbf{x}}} = n\,\dot{\mathbf{x}}\,.$$

This gives for the Euler–Lagrange equation

$$\frac{d}{ds}\frac{\partial L}{\partial \dot{\mathbf{x}}} - \frac{\partial L}{\partial \mathbf{x}} = 0$$

the well known ray equation

$$\frac{d}{ds}\left(n\,\frac{d\mathbf{x}}{ds}\right) = \boldsymbol{\nabla} n. \tag{1.17}$$

We give another derivation of (1.11) by making use of the eikonal description. For the metric (1.13) the eikonal equation $g^{\mu\nu}\partial_\mu S\partial_\nu S = 0$ becomes, by setting $S(x) = \hat{S}(\mathbf{x}) - \omega t$,

$$\left(\boldsymbol{\nabla}\hat{S}\right)^2 = n^2\omega^2, \quad n = 1 - \frac{2U}{c^2}, \tag{1.18}$$

where the operations on the left have to be understood in the Euclidean 3-space.

Alternatively, one can write the Maxwell equations for the metric g as in Special Relativity, but with

$$\varepsilon = \mu \simeq 1 - \frac{2U}{c^2}. \qquad \textbf{(exercise)} \tag{1.19}$$

Hence, Maxwell's relation gives again

$$n = \sqrt{\varepsilon\mu} \simeq 1 - \frac{2U}{c^2}. \tag{1.20}$$

Solution: Maxwell eqs. in an almost Newtonian g-field

We decompose F into electric and magnetic pieces, $F = E \wedge dt + B$ The homogeneous equation $dF = 0$ splits into

$$dB = 0, \quad dE + \partial_t B = 0. \tag{1.21}$$

If we set $*F = D - H \wedge dt$, we have similarly from $d * F = 0$:

$$dD = 0, \quad dH = \partial_t D. \tag{1.22}$$

For a static metric ${}^{(4)}g = -\alpha^2 dt^2 + g$ we obtain the relations

$$D = \frac{1}{\alpha} * E, \quad B = \frac{1}{\alpha} * H. \tag{1.23}$$

For an almost Newtonian situation: $g \simeq \alpha^{-2}\delta$ (δ: flat metric). Thus

$$*E = \frac{1}{\alpha}\bar{*}E \quad (\bar{*}: \text{ flat star operator})$$

and hence

$$D = \frac{1}{\alpha^2}\bar{*}E, \quad B = \frac{1}{\alpha^2}\bar{*}H. \tag{1.24}$$

(1.21), (1.22) and (1.24) are just Maxwell's equations in flat space-time for a *medium* with

$$\varepsilon = \mu \simeq \frac{1}{\alpha^2} \to n = \sqrt{\varepsilon\mu} \simeq \frac{1}{\alpha^2} = 1 - 2\frac{U}{c^2}. \tag{1.25}$$

1.2 Deflection by an arbitrary mass concentration

In terms of the unit tangent vector $\mathbf{e} = d\mathbf{x}/ds$, Eq. (1.17) can be approximated as

$$\frac{d}{ds}\mathbf{e} = -\frac{2}{c^2}\nabla_\perp U, \tag{1.26}$$

where $\nabla_\perp$ denotes the transverse derivative, $\nabla_\perp = \nabla - \mathbf{e}(\mathbf{e}\cdot\nabla)$. This gives for the deflection angle $\hat{\boldsymbol{\alpha}} = \mathbf{e}_{in} - \mathbf{e}_{fin}$, with initial and final directions $\mathbf{e}_{in}$ and $\mathbf{e}_{fin}$, respectively,

$$\hat{\boldsymbol{\alpha}} = \frac{2}{c^2}\int_{u.p.} \nabla_\perp U ds, \tag{1.27}$$

where the integral is taken over the unperturbed path (*u.p.*).[1] Here, we insert the expression (1.12) for the Newtonian potential of a mass density $\rho(\mathbf{x})$. Parametrizing the unperturbed path as $\mathbf{x} = \boldsymbol{\xi} + z\mathbf{e}$, $\mathbf{e} \equiv \mathbf{e}_{in}$, $\boldsymbol{\xi}$ fixed, we obtain

$$\hat{\boldsymbol{\alpha}} = \frac{2}{c^2}\int \nabla_\perp U(\boldsymbol{\xi} + z\mathbf{e})\, dz = -\frac{2}{c^2}G\int d^3x'\rho(\mathbf{x}')\int \nabla_\perp\left(\frac{1}{|\mathbf{x}-\mathbf{x}'|}\right) dz$$

or, setting $\mathbf{x}' \equiv (\boldsymbol{\xi}', z')$ (*i.e.*, $\mathbf{x}' = \boldsymbol{\xi}' + z'\mathbf{e}$) and using

$$\nabla_\perp\left(\frac{1}{|\mathbf{x}-\mathbf{x}'|}\right) = -\frac{\boldsymbol{\xi}-\boldsymbol{\xi}'}{|\mathbf{x}-\mathbf{x}'|^3},$$

[1] Alternatively we can proceed from here as follows. From (1.27) we obtain

$$\nabla_\perp \cdot \hat{\boldsymbol{\alpha}} = \frac{2}{c^2}\int_{u.p.} \Delta_\perp U\, ds. \tag{1.28}$$

Here, we can replace the transversal Laplacian by the three-dimensional one and use the Poisson equation $\Delta U = 4\pi G\rho$:

$$\nabla_\perp \cdot \hat{\boldsymbol{\alpha}} = \frac{8\pi G}{c^2}\int_{u.p.} \rho\, ds = \frac{8\pi G}{c^2}\Sigma. \tag{1.29}$$

On the other hand, (1.27) can be written as

$$\hat{\boldsymbol{\alpha}} = \frac{2}{c^2}\nabla_\perp\hat{\psi}, \quad \hat{\psi} = \int_{u.p.} U\, ds. \tag{1.30}$$

Hence, we have

$$\Delta_\perp\hat{\psi} = 4\pi G\Sigma. \tag{1.31}$$

Using the Green's function (1.37) below, the potential $\hat{\psi}$ is given by

$$\hat{\psi}(\boldsymbol{\xi}) = 2G\int \ln|\boldsymbol{\xi}-\boldsymbol{\xi}'|\,\Sigma(\boldsymbol{\xi}')d^2\boldsymbol{\xi}'. \tag{1.32}$$

From this we obtain for the deflection angle

$$\hat{\boldsymbol{\alpha}} = \frac{4G}{c^2}\int_{\mathbb{R}^2} \frac{\boldsymbol{\xi}-\boldsymbol{\xi}'}{|\boldsymbol{\xi}-\boldsymbol{\xi}'|}\Sigma(\boldsymbol{\xi}')d^2\boldsymbol{\xi}', \tag{1.33}$$

which agrees with (1.34).

we have

$$\hat{\alpha} = \frac{2G}{c^2} \int d^2\xi' dz' \rho\left(\xi', z'\right)\left(\xi - \xi'\right) \int \frac{dz}{\left[\left(\xi - \xi'\right)^2 + \left(z - z'\right)^2\right]^{3/2}}.$$

The z–integration has to be extended over the interval between the source and the observer. Now, since the extension of the lens (for instance a cluster of galaxies) is much smaller than the distances of the observer and the source to the lens, one is allowed to replace the finite interval for z by all of $\mathbb{R}$. Making use of

$$\int_{-\infty}^{+\infty} \frac{dz}{\left[\left(\xi - \xi'\right)^2 + \left(z - z'\right)^2\right]^{3/2}} = \frac{2}{\left|\xi - \xi'\right|^2},$$

$\left(\text{note that } \frac{1}{\mathrm{a}^2} \frac{\partial}{\partial \mathrm{u}} \left(\frac{\mathrm{u}}{\sqrt{\mathrm{a}^2 + \mathrm{u}^2}}\right) = \frac{1}{[\mathrm{a}^2 + \mathrm{u}^2]^{3/2}}\right)$, we finally get

$$\hat{\alpha}\left(\xi\right) = \frac{4G}{c^2} \int d^2\xi' \frac{\xi - \xi'}{\left|\xi - \xi'\right|^2} \int dz' \rho\left(\xi', z'\right).$$

Only the *projected mass density*

$$\Sigma\left(\xi\right) = \int \rho\left(\xi, z\right) dz,$$

appears on the right hand side. In terms of this, our main result of this section becomes

$$\hat{\alpha}\left(\xi\right) = \frac{4G}{c^2} \int_{\mathbb{R}^2} \frac{\xi - \xi'}{\left|\xi - \xi'\right|^2} \Sigma\left(\xi'\right) d^2\xi'. \tag{1.34}$$

For a point mass M located at the origin of the transversal plane, $\Sigma\left(\xi\right)$ is equal to $M\delta^2\left(\xi\right)$ and thus

$$\hat{\alpha}\left(\xi\right) = \frac{4GM}{c^2} \frac{1}{|\xi|}, \tag{1.35}$$

which is Einstein's famous result. Consider, more generally, an axially symmetric lens with mass $M\left(\xi\right)$ located inside the cylinder defined by the impact parameter ξ, we expect from (1.35) that

$$\hat{\alpha}\left(\xi\right) = \frac{4G}{c^2} \frac{M\left(\xi\right)}{\xi}. \tag{1.36}$$

For a simple way to show this, we note first that

$$\mathcal{G}\left(\xi\right) = \frac{1}{2\pi} \ln |\xi| \tag{1.37}$$

is the Green's function of the 2–dimensional Laplacian[2]

$$^{(2)}\Delta\mathcal{G} = \delta^{(2)}. \tag{1.38}$$

[2]Proof of $\Delta\mathcal{G} = \delta$:

It is easy to verify that for $|\mathbf{x}| > 0$ $\Delta\mathcal{G} = 0$ holds: in polar coordinates (r, φ) we have

$$\Delta \ln |\mathbf{x}| = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \ln r}{\partial r}\right) = 0 \qquad \text{for} \qquad \mathrm{r} > 0.$$

Let f be a test function with $\operatorname{supp} f \subset D_R$ (disk of radius R). Then

$$< \Delta \ln |\mathbf{x}|, f >=< \ln |\mathbf{x}|, \Delta f >= \lim_{\varepsilon \to 0} \int_{\varepsilon < |\mathbf{x}| < R} \ln |\mathbf{x}| \Delta f d^2x.$$

For the last integral the second Green's formula gives

$$\int_{\varepsilon < |\mathbf{x}| < R} \Delta \ln |\mathbf{x}| f d^2x + \left(\int_{|\mathbf{x}|=\varepsilon} + \int_{|\mathbf{x}|=R}\right) \left(\ln |\mathbf{x}| \frac{\partial f}{\partial n} - f \frac{\partial \ln |\mathbf{x}|}{\partial n}\right) ds =$$

$$\int_{|\mathbf{x}|=\varepsilon} \left(-\ln |\mathbf{x}| \frac{\partial f}{\partial r} + f \frac{1}{r}\right) ds.$$

Here, only the second term survives in the limit $\varepsilon \to 0$ and becomes equal to $2\pi f(0)$. This shows that $< \Delta \ln |\mathbf{x}|, f >= 2\pi < \delta, f >$ for all test functions f, which proves our claim.

With its help the result (1.34) can be written briefly as the gradient

$$\hat{\alpha} = \frac{8\pi G}{c^2} \nabla (\mathcal{G} * \Sigma), \tag{1.39}$$

where the star ($*$) denotes convolution. Taking the divergence of this equation gives, with (1.38),

$$\nabla \cdot \hat{\alpha} = 8\pi G \Sigma / c^2. \tag{1.40}$$

We integrate this over the disk with radius ξ and obtain with the 2–dimensional version of Gauss' theorem for an axially symmetric lens

$$\int \nabla \cdot \hat{\alpha}\, d^2\xi = \int \hat{\alpha} \cdot \mathbf{n}\, ds = 2\pi\xi\hat{\alpha}(\xi) = 8\pi G M(\xi)$$

($\mathbf{n}$: unit outerward normal to the disk). The last equality just gives (1.36). (You also can obtain this result by introducing polar coordinates in (1.34) and working out the angular integration, but the above derivation is much simpler.)

We shall study axisymmetric lens models in detail later, but I want to proceed now with the general lens.

1.3 The general lens map

Fig. 1.2 summarizes some of the notation we are using. (I will follow as much as possible the beautiful monograph [5] by Schneider, Ehlers and Falco, hereafter quoted as SEF.) Simple geometry shows that

$$\boldsymbol{\eta} = \frac{D_s}{D_d}\boldsymbol{\xi} - D_{ds}\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}), \tag{1.41}$$

where $\boldsymbol{\eta}$ is the source position and $\boldsymbol{\xi}$ is the lens position. This defines a map $\boldsymbol{\xi} \to \boldsymbol{\eta}$ from the lens plane to the source plane. Fig. 1.2 shows also that

$$\boldsymbol{\xi} = D_d\boldsymbol{\theta}, \qquad \boldsymbol{\eta} = D_s\boldsymbol{\beta}, \tag{1.42}$$

hence (1.41) can be written as

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{ds}}{D_s}\hat{\boldsymbol{\alpha}}. \tag{1.43}$$

This or (1.41) is what is called the *lens equation.*[3]

It turns out, not unexpectedly, that (1.43) holds also in cosmology (see Section 4.1).

It is convenient to write (1.41) in dimensionless form. Let ξ_0 be a length parameter in the lens plane (whose choice will depend on the specific problem), and let η_0 be the corresponding length in the source plane, $\eta_0 = (D_s/D_d)\,\xi_0$. We set $\mathbf{x} = \boldsymbol{\xi}/\xi_0$, $\mathbf{y} = \boldsymbol{\eta}/\eta_0$, and

$$\kappa(\mathbf{x}) = \frac{\Sigma(\xi_0\mathbf{x})}{\Sigma_{crit}}, \qquad \boldsymbol{\alpha}(\mathbf{x}) = \frac{D_d D_{ds}}{\xi_0 D_s}\hat{\boldsymbol{\alpha}}(\xi_0\mathbf{x}), \tag{1.44}$$

with

$$\Sigma_{crit} = \frac{c^2}{4\pi G} \cdot \frac{D_s}{D_d D_{ds}} = 0.35 \text{ g cm}^{-2}\left(\frac{1 \text{ Gpc}}{D_d D_{ds}/D_s}\right). \tag{1.45}$$

Then Eq. (1.41) reads as follows

$$\mathbf{y} = \mathbf{x} - \boldsymbol{\alpha}(\mathbf{x}), \tag{1.46}$$

whereby Eq. (1.34) translates to

$$\boldsymbol{\alpha}(\mathbf{x}) = \frac{1}{\pi}\int_{\mathbb{R}^2} \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|^2}\kappa(\mathbf{x}')\, d^2x'. \tag{1.47}$$

[3]See the appendix 1.8 for an alternative derivation.

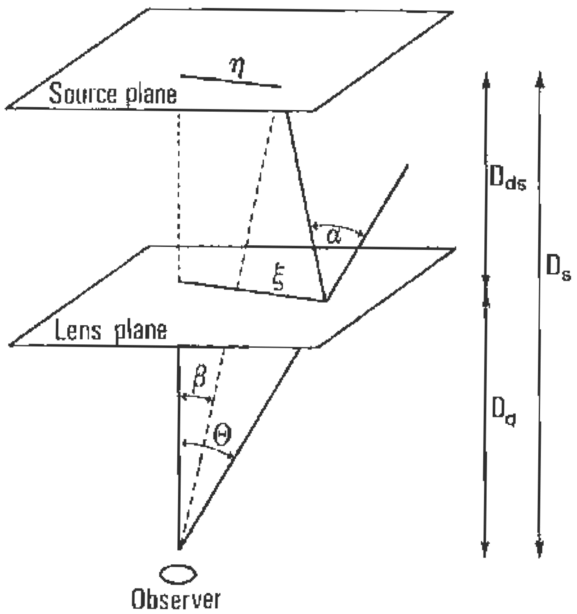

Figure 1.2: Notation adopted for the description of the lens geometry.

As in (1.39) we can write $\boldsymbol{\alpha}$ as a gradient

$$\boldsymbol{\alpha} = \boldsymbol{\nabla}\psi, \quad \psi = 2\mathcal{G} * \kappa. \tag{1.48}$$

From (1.38) it follows that ψ satisfies the 2–dimensional Poisson equation

$$\Delta\psi = 2\kappa. \tag{1.49}$$

The map $\varphi : \mathbf{x} \mapsto \mathbf{y}$ defined by (1.46) is thus a gradient map

$$\varphi(\mathbf{x}) = \boldsymbol{\nabla}\left(\frac{1}{2}\mathbf{x}^2 - \psi(\mathbf{x})\right). \tag{1.50}$$

Explicitly, the second equation in (1.48) reads

$$\psi(\mathbf{x}) = \frac{1}{\pi}\int_{\mathbb{R}^2} \ln(|\mathbf{x}-\mathbf{x}'|)\,\kappa(\mathbf{x}')\,d^2x'. \tag{1.51}$$

The differential $D\varphi$ will often be used. A standard parametrization is

$$D\varphi = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}, \tag{1.52}$$

where $\gamma_1 = \frac{1}{2}(\partial_{11}\psi - \partial_{22}\psi)$, $\gamma_2 = \partial_{12}\psi = \partial_{21}\psi$. Note, that the matrix elements of $D\varphi$ are (see, e.g., (1.50))

$$(D\varphi)_{ij} = \delta_{ij} - \partial_i\partial_j\psi, \tag{1.53}$$

where $\partial_i = \partial/\partial x_i$. In particular, the trace of (1.52) is chosen correctly (see (1.49)), and $D\varphi$ is clearly symmetric. The dimensionless projected mass density κ is often called the *mean Ricci curvature*, and the 2–dimensional vector $\boldsymbol{\gamma} = (\gamma_1, \gamma_2)$ is the *Weyl shear*. For a geometrical interpretation, see Section 1.7.

1.4 Magnification, critical curves and caustics

Next, I show that the magnification μ, that is, the ratio of the flux of an image to the flux of the unlensed source, is given by

$$\mu = \frac{1}{|\det D\varphi|}. \tag{1.54}$$

In order to derive this result, I recall a simple but important fact from ray optics.

Consider a ray L and any two points along the ray and construct areas dA_1 and dA_2 normal to the ray at these points. Let $dE_1 = dE_2$ be the energy of all rays passing through both dA_1 and dA_2 during the time dt. Since

$$dE_1 = I_{\nu_1} dA_1 dt d\Omega_1 d\nu_1, \qquad dE_2 = I_{\nu_2} dA_2 dt d\Omega_2 d\nu_2,$$

where $d\Omega_{1,2}$ is the solid angle subtended by $dA_{2,1}$ at $dA_{1,2}$, and because $d\Omega_1 = dA_2/R^2$, $d\Omega_2 = dA_1/R^2$ ($R =$ distance between dA_1 and dA_2), we obtain

$$I_{\nu_1} d\nu_1 = I_{\nu_2} d\nu_2. \tag{1.55}$$

If there is no frequency shift, the specific intensity is thus constant along a ray

$$I_\nu = const. \tag{1.56}$$

This holds also for gravitational light deflection by localized, nearly static lenses, because this does not introduce an additional frequency shift between source and observer, beside the cosmological one (cosmological effects will be considered later).[4] Now, the flux of an image of an infinitesimal source is the product of its surface brightness I and the solid angle $d\Omega$ it subtends on the sky. From (1.56) we conclude that the magnification μ is the ratio of $d\Omega$ and the solid angle $d\Omega_0$ for the undeflected situation. On the other hand $d\Omega_0/d\Omega$ is equal to the area distortion of the lens map φ, and thus equal to the Jacobian $|\det\ D\varphi|$. This proves our claim (1.54). (For a more sophisticated derivation, which applies also in cosmology, see SEF, Sections 3.4–3.6, in particular eqs. (3.81), (3.82).)

The lens map φ becomes singular along *critical curves* in the lens plane. These are characterized by

$$\det(D\varphi) = 0 \tag{1.57}$$

or

$$(1-\kappa)^2 - |\gamma|^2 = 0. \tag{1.58}$$

The *caustics* are the images of these critical curves[5].

In the vicinity of these source points the magnification becomes very large. On caustics it diverges formally, but this is of no physical significance, because the magnification remains finite for any extended source (see Section 2.6 or SEF, Section 6.4). For a point–like source, the ray approximation breaks down and we would have to use wave optics. (You find a discussion of this in Chapter 7 of SEF.)

1.5 Time delay

The lens map (1.50) can also be written as

$$\nabla_{\mathbf{x}} \phi(\mathbf{x}, \mathbf{y}) = 0, \tag{1.59}$$

[4]More general argument: Without absorption and scattering processes the distribution function f for photons satisfies the Liouville equation

$$L_{X_g} f = 0,$$

where X_g is the vector field of the geodesic spray. Since the intesity $I(\omega)$ is proportional to $\omega^3 f$, we conclude that $I(\omega)/\omega^3$ *remains constant along null-geodesics.*

[5]A famous theorem of Sard tells us that these critical values of φ form a set of measure zero.

with

$$\phi(\mathbf{x},\mathbf{y}) = \frac{1}{2}(\mathbf{x}-\mathbf{y})^2 - \psi. \tag{1.60}$$

This formulation reflects the Fermat principle. We now show that the delay of arrival times is directly given by the *Fermat potential* ϕ

$$c\Delta t = \xi_0^2 \frac{D_s}{D_d D_{ds}} \phi(\mathbf{x},\mathbf{y}). \tag{1.61}$$

The travel time is $1/c$ times the integral in the Fermat principle (1.15)

$$\begin{aligned} t &= \frac{1}{c}\int_{path} n(\mathbf{x}(s))\,|\dot{\mathbf{x}}(s)|\,ds \simeq \\ &\simeq \frac{l}{c} - \frac{2}{c^3}\int_{u.p.} U(\mathbf{x}(s))\,ds \end{aligned} \tag{1.62}$$

(l = path length). In (1.63) we have used the basic formula (1.11) for the effective refraction index. It suffices to take the integral along the unperturbed path. (This part describes the "Shapiro time delay".)

A look at Fig. 1.2 shows that

$$l = \sqrt{D_{ds}^2 + (\boldsymbol{\xi}-\boldsymbol{\eta})^2} + \sqrt{\boldsymbol{\xi}^2 + D_d^2} \simeq D_{ds} + D_d + \frac{1}{2D_{ds}}(\boldsymbol{\xi}-\boldsymbol{\eta})^2 + \frac{1}{2D_d}\boldsymbol{\xi}^2.$$

For the potential part in (1.63) we proceed as in Section 1.2 (using the same notation). We have

$$\begin{aligned} \int_{u.p.} U(\mathbf{x}(s))\,ds &= \int U(\boldsymbol{\xi} + z\mathbf{e})\,dz = \\ &= -G\int d^2\xi' dz' \rho(\xi', z') \int \frac{dz}{\sqrt{(\boldsymbol{\xi}-\boldsymbol{\xi}')^2 + (z-z')^2}}. \end{aligned} \tag{1.63}$$

Now z' varies over a much smaller domain than z. Therefore, the z–integral is approximately equal to

$$\begin{aligned} \int_{-D_d}^{+D_{ds}} \frac{du}{\sqrt{u^2 + (\boldsymbol{\xi}-\boldsymbol{\xi}')^2}} &= \ln(u + \sqrt{u^2 + (\boldsymbol{\xi}-\boldsymbol{\xi}')^2}\,)\Big|_{-D_d}^{+D_{ds}} \simeq \\ &\simeq \ln\left[\frac{2D_{ds}}{|\boldsymbol{\xi}-\boldsymbol{\xi}'|}\right] + \ln\left[\frac{2D_d}{|\boldsymbol{\xi}-\boldsymbol{\xi}'|}\right]. \end{aligned} \tag{1.64}$$

Since the last expression is independent of z', we can do the z'–integral and find

$$-\frac{2}{c^3}\int U(\mathbf{x}(s))\,ds = -\frac{4G}{c^3}\int d^2\xi' \Sigma(\boldsymbol{\xi}') \ln\left[\frac{|\boldsymbol{\xi}-\boldsymbol{\xi}'|}{\xi_0}\right] + const. \tag{1.65}$$

If we subtract from (1.63) the arrival time for an unlensed ray from the same source, we obtain the time delay

$$\Delta t = \frac{1}{c}\frac{D_d D_s}{2D_{ds}}\left(\frac{\boldsymbol{\xi}}{D_d} - \frac{\boldsymbol{\eta}}{D_s}\right)^2 - \frac{1}{c}\hat{\psi}(\boldsymbol{\xi}) + const, \tag{1.66}$$

where

$$\hat{\psi}(\boldsymbol{\xi}) = \frac{4G}{c^2}\int d^2\boldsymbol{\xi}' \Sigma(\boldsymbol{\xi}') \ln\left[\frac{|\boldsymbol{\xi}-\boldsymbol{\xi}'|}{\xi_0}\right]. \tag{1.67}$$

The constant in (1.65) is independent of $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$. From the definitions in Section 1.3 one finds that

$$\hat{\psi}(\boldsymbol{\xi}) = \frac{D_s \xi_0^2}{D_d D_{ds}} \psi(\mathbf{x}) \qquad (\boldsymbol{\xi} = \xi_0 \mathbf{x}), \tag{1.68}$$

and in terms of the dimensionless variables $\mathbf{x}$ and $\mathbf{y}$ one finds that (1.65) agrees with the claimed formula (1.61).

In a cosmological context one has to take into account red–shift effects (see Sect. 4.1). It should, however, already be clear that the prefactor in (1.60) involves the Hubble parameter.

1.6 Whitney theorem on generic singularities

In a pioneering work of modern singularity theory, H. Whitney [6] studied in 1955 the generic properties of smooth mappings of the plane into itself. His results apply directly to realistic lens maps, and it is, therefore, interesting to know what he showed. Before telling you, I must explain what fold and cusp singularities are. Consider a smooth map $f : \mathbb{R}^2 \to \mathbb{R}^2$. We say that $\mathbf{x} \in \mathbb{R}^2$ is a

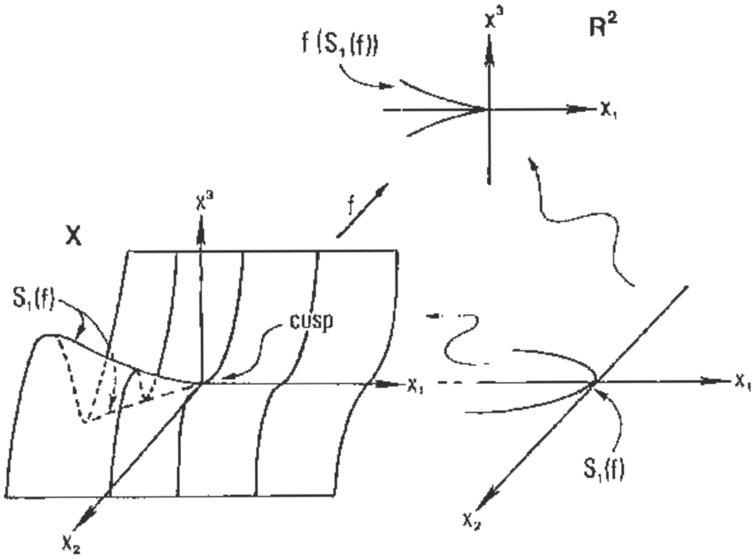

Figure 1.3: Cusp singularity. The map from the plane $(x_1,\ x_2)$ to X is given by $(x_1,\ x_2) \mapsto (x_1,\ x_2,\ x_1x_2 + x_2^3)$; $S_1(f)$ in coordinates is given by $x_1 = -3x_2^2$. The map from the plane $(x_1,\ x_2)$ to the plane $(x_1,\ x_3)$ is given by $(x_1,\ x_2) \mapsto (x_1,\ x_1x_2 + x_2^3)$, and corresponds to the normal form for a cusp point; $f(S_1(f))$ is given as $t \to (-3t^2,\ -2t^3)$.

regular point, if the differential $Df\,(\mathbf{x})$ is non–singular, *i.e.*, $\det Df \neq 0$ at $\mathbf{x}$. In this situation the map is locally differentially equivalent to the identity $(x_1, x_2) \longmapsto (x_1, x_2)$ in a sufficiently small neighborhood of the origin. (Technically, it is convenient to use germs of maps.)

If $Df\,(\mathbf{x})$ is singular at x, one calls $\mathbf{x}$ a *critical point* and its image $f(\mathbf{x})$ a *critical value.* We say that a critical point is a *fold* if the germ of the map at that point is diffeomorphic to the germ

$$(x_1, x_2) \longmapsto (x_1^2, x_2) \qquad at \quad x_1 = x_2 = 0. \tag{1.69}$$

This means that one can introduce local coordinates (x_1, x_2) in the source plane and (y_1, y_2) in the target plane, such that locally $y_1 = x_1^2,\ y_2 = x_2$.

We say that a critical point is a *cusp*, if the map is locally equivalent to

$$(x_1, x_2) \to \left(x_1^3 + x_1x_2, x_2\right). \tag{1.70}$$

An example of a cusp is shown in Fig. 1.4. Whitney has proven that maps which have only fold and cusp singularities are *generic.* (This means that these contain an open and dense subset of all smooth maps in some natural topology, now called the Whitney topology.) Moreover, those maps in this open and dense subset which satisfy a few mild global conditions are also *stable.* (I do not want to give a precise definition for stability.) In higher dimension, things are much more complicated.

1.7 Classification of ordinary images, orientation and shape

For a given source position $\mathbf{y}$ the images are those points $\mathbf{x}$ which satisfy $\nabla\phi(\mathbf{x}) = 0$ (critical points of the Fermat potential). This invites us to use some standard mathematical concepts and tools for a qualitative discussion of the lens map.

A critical point of ϕ is *non–degenerate*, if the Hessian, $H(\phi) = (\phi_{,ij}) = D\varphi$, is a non–degenerate quadratic form (equivalently: det $D\varphi \neq 0$), that is, the source is not on a caustic. The *index* of such a critical point is just the index of $H(\phi)$ at that point, *i.e.*, equal to the number of negative terms in the normal form of $H(f)$. In two dimensions we have three types of non–degenerate critical points $\equiv$ *ordinary* images:

- *type I* corresponds to a minimum of ϕ (index = 0),
- *type II* to a saddle (index = 1),
- *type III* to a maximum (index = 2).

For a given source position, not lying on a caustic, we denote by n_I, n_{II}, n_{III} the number of ordinary images of the indicated type. Using a theorem by Morse, one can show that

$$n_I - n_{II} + n_{III} = 1. \tag{1.71}$$

According to this theorem, the left–hand side of (1.70) is equal to the Euler characteristic of a big circle, which is equal to 1. In Chapter 5, I will give an elementary derivation of this result by making use of standard tools in complex analysis.

As a consequence of (1.70) we arrive at the important fact

$$n := n_I + n_{II} + n_{III} = 1 + 2n_{II}. \tag{1.72}$$

The number of ordinary images of a regular lens is thus always odd. This number is ***bigger than one***, if and only if the ***arrival time surface*** $\{\phi = const\}$ ***has saddle points.*** A beautiful example is shown in Figs. 1.4, 1.5, taken from [7]. Remember: ordinary images are located at local extrema and a saddle points of the arrival time surface. We say that such an image has ***positive*** (***negative***) ***parity*** if det$D\varphi > 0 (< 0)$. Clearly, type I and III have positive parity, while type II has negative parity. There is a simple geometrical meaning of this notion. Consider a source at $\mathbf{y}$ and an image at $\mathbf{x}$, $\mathbf{y} = \varphi(\mathbf{x})$. The linearization $A := D\varphi(\mathbf{x})$ of φ at $\mathbf{x}$ tells us what happens with infinitesimal displacements, described by vectors $\mathbf{Y}$ at the source and vectors $\mathbf{X}$ at the image point $\mathbf{x}$, related by $\mathbf{Y} = A\mathbf{X}$.

Note first, that $\mathbf{X}\cdot\mathbf{Y} = A(\mathbf{X},\mathbf{X}) := \Sigma A_{ij}X_iX_j > 0 (< 0)$ for type I, III, (type II). Thus, for images of type I, III the position angle of the image vector differs by no more than $\pi/2$ from that of the source, while for type II they differ by more than $\pi/2$. Consider now two pairs $\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}$ and $\mathbf{X}^{(2)}, \mathbf{Y}^{(2)}$, with $\mathbf{Y}^{(i)} = A\mathbf{X}^{(i)}$. Let $\mathbf{X}^{(1)} \wedge \mathbf{X}^{(2)} = \det\left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}\right)$, etc. Then $\mathbf{Y}^{(1)} \wedge \mathbf{Y}^{(2)} = (\det A)\mathbf{X}^{(1)} \wedge \mathbf{X}^{(2)}$. This shows that positive parity images ***preserve the handedness,*** whereas for negative parity (type II) it is ***reversed.***

Finally, let us see how infinitesimal circles in the source plane are deformed. Thus, consider a small circular source with radius R at $\mathbf{y}$, bounded by the curve

$$\mathbf{c}(t) = \mathbf{y} + \begin{pmatrix} R\,\cos t \\ R\,\sin t \end{pmatrix} \qquad (0 \leq t \leq \pi).$$

The boundary of the images is

$$\mathbf{d}(t) = \mathbf{x} + A^{-1}\begin{pmatrix} R\,\cos t \\ R\,\sin t \end{pmatrix}.$$

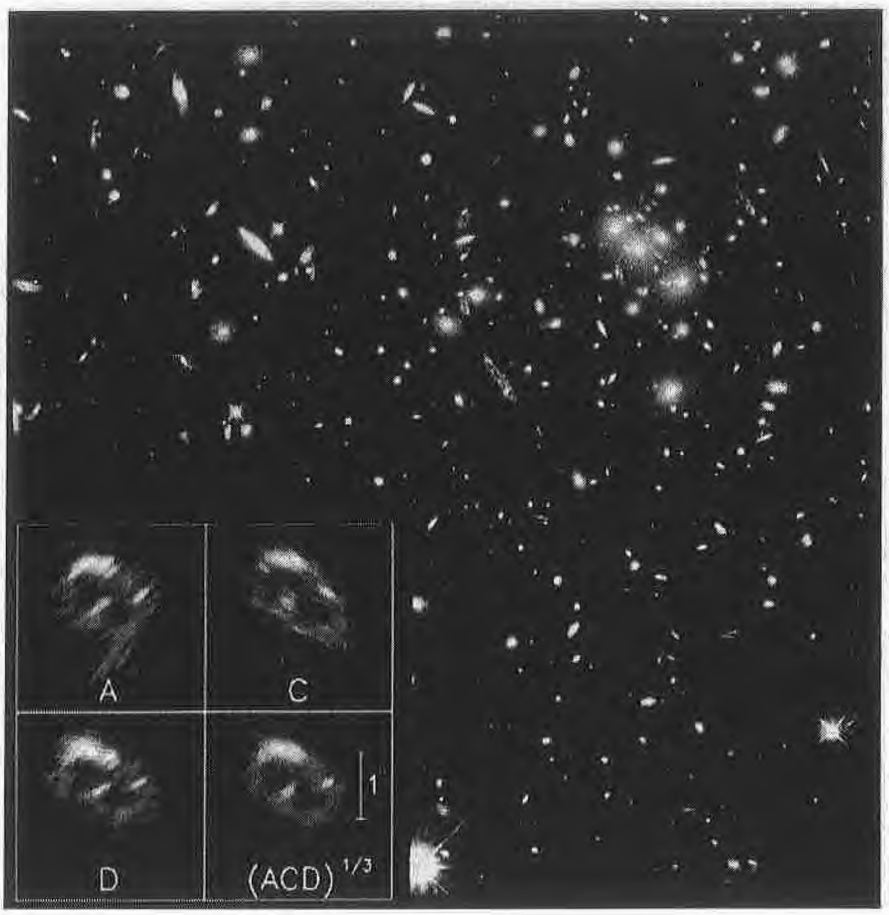

Figure 1.4: Five blue arcs, each an image of the same background source at high redshift, lensed by the massive cluster Cl0024+1654 at $z = 0.4$ (taken from W. Calley, T. Tyson and E. Turner, 1996 [7]).

Inserting the parametrization (1.52) one finds that this describes an ellipse with semi–axis parallel to the principle axis of A, with magnitudes

$$\frac{R}{|1-\kappa \pm |\gamma||}, \tag{1.73}$$

and position angles $\varphi_\pm$ for the axis given by

$$\tan 2\varphi_\pm = -\gamma_2/\gamma_1. \tag{1.74}$$

We leave it as an exercise to derive these results. *Hints:* Write $\mathbf{d}\,(t)$ in complex form. One finds that

$$\begin{pmatrix} X \\ Y \end{pmatrix} := A^{-1} \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \tag{1.75}$$

can be written in the form (familiar from optics)

$$X \pm iY = \frac{1}{2}\left(E_\pm e^{-it} + E_\pm^* e^{it}\right), \tag{1.76}$$

with

$$E_+ = \frac{\gamma}{(1-\kappa)^2 - |\gamma|^2}, \qquad E_- = \frac{1-\kappa}{(1-\kappa)^2 - |\gamma|^2} \qquad (\gamma = \gamma_1 + i\gamma_2). \tag{1.77}$$

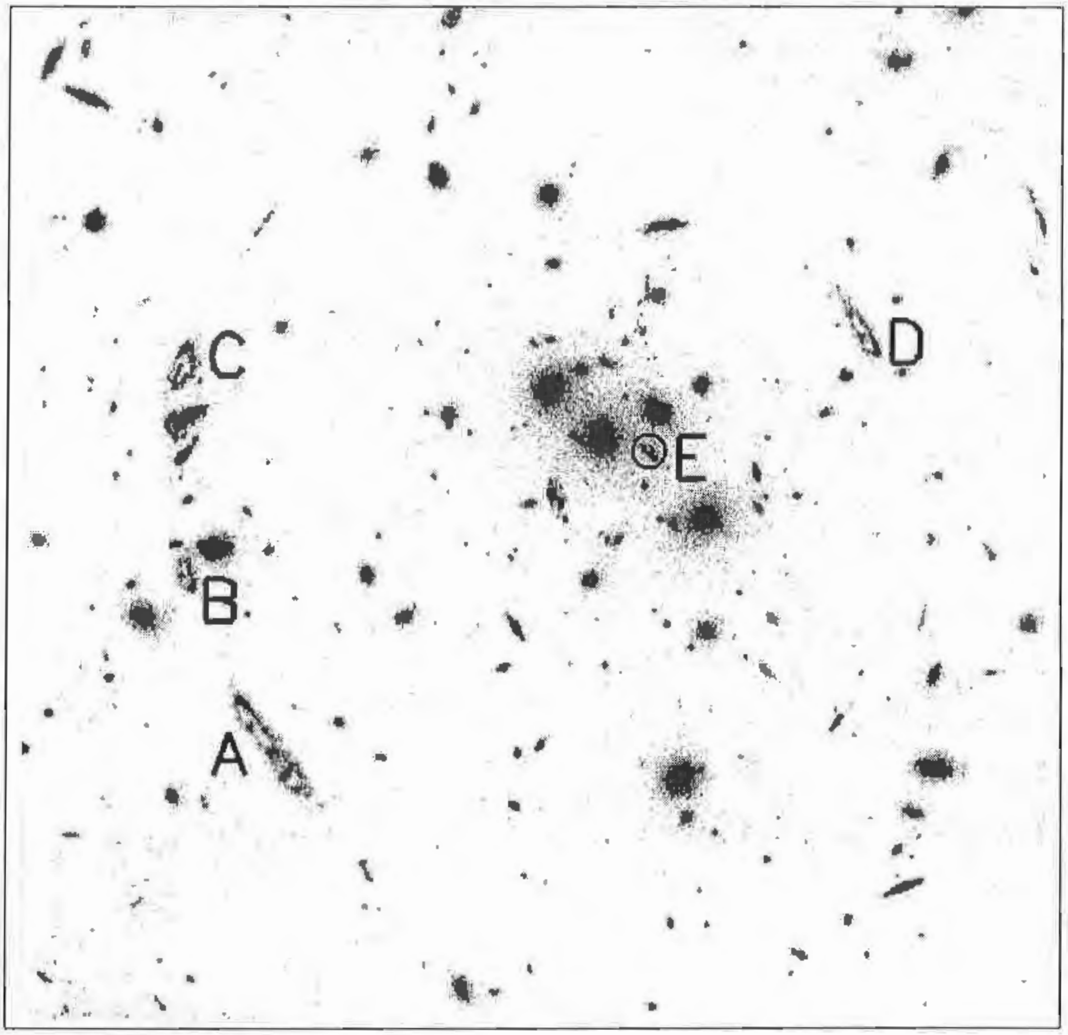

Figure 1.5: The five images (arcs) of the same blue source galaxy in Fig. 1.4.

The ratio E_+/E_- describes the orientation and shape of the ellipse, and is thus observable. This complex number is what is called the complex *reduced shear* (P. Schneider):

$$g = \frac{\gamma}{1-\kappa} \qquad \left(\mathrm{g} = \frac{\gamma}{1-\kappa}\right). \tag{1.78}$$

1.8 Appendix: *Alternative derivation of the lens equation*

For the later generalization of the lens equation (1.43) to a cosmological situation (Friedmann background), I give here another derivation.

By making use of the notation in the figure 1.6, the differential equation for a light ray $\boldsymbol{\xi} = \boldsymbol{\xi}(\boldsymbol{\theta}, \zeta)$ ($\boldsymbol{\theta}$: observing angles) is, to first order in the Newtonian potential ($c \to 1$):

$$\frac{d^2\boldsymbol{\xi}}{d\zeta^2} = -2\boldsymbol{\nabla}_\perp U(\boldsymbol{\xi}\,(\boldsymbol{\theta}, \zeta), \zeta)\,. \tag{1.79}$$

(Note that $d\zeta/ds = 1$, up to quadratic terms in the small ξ_i.)

With the help of the Green's function $G(\zeta, \zeta')$ of the operator $d^2/d\zeta^2$, given by

$$G(\zeta, \zeta') = (\zeta - \zeta')\theta(\zeta - \zeta'), \tag{1.80}$$

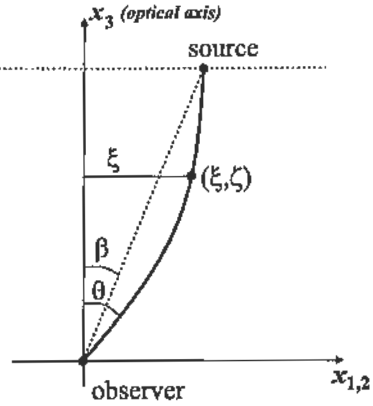

Figure 1.6: Notation adopted for the description of the lens geometry.

we transform (1.79) into an integral equation,

$$\boldsymbol{\xi}\left(\boldsymbol{\theta},\zeta\right),\zeta) = \zeta\boldsymbol{\theta} - 2\int_0^{\zeta} d\zeta'(\zeta-\zeta')\nabla_{\perp}U(\boldsymbol{\xi}\left(\boldsymbol{\theta},\zeta\right),\zeta)\,. \tag{1.81}$$

Here, the first term is a solution of the homogeneous equation, which would describe the light ray without deflection. Clearly,

$$\frac{d}{d\zeta}\boldsymbol{\xi}(\boldsymbol{\theta},\zeta)\bigg|_{\zeta=0} = \boldsymbol{\theta}. \tag{1.82}$$

To first order in U, we can replace the argument in $\nabla_{\perp}U$ on the right hand side of (1.81) by the unperturbed ray and obtain the explicit solution

$$\boldsymbol{\xi}(\boldsymbol{\theta},\zeta) = \zeta\boldsymbol{\theta} - 2\int_0^{\zeta} d\zeta'\,(\zeta-\zeta')\,\nabla_{\perp}U(\zeta'\boldsymbol{\theta},\zeta'). \tag{1.83}$$

At the source plane ($\zeta = \zeta_s$) this gives, with $\boldsymbol{\eta} \equiv \boldsymbol{\xi}(\boldsymbol{\theta},\zeta_s)$,

$$\boldsymbol{\eta} = \zeta_S\boldsymbol{\theta} - 2\int_0^{\zeta_s} d\zeta(\zeta_S-\zeta)\nabla_{\perp}U(\zeta\boldsymbol{\theta},\zeta). \tag{1.84}$$

Since $\boldsymbol{\eta} = \zeta_s\boldsymbol{\beta}$, we obtain the lens equation

$$\boldsymbol{\beta} = \boldsymbol{\theta} - 2\int_0^{\zeta_s} d\zeta\,\underbrace{\frac{\zeta_s-\zeta}{\zeta_s}}_{(1)}\,\nabla_{\perp}U\,\underbrace{(u.p.)}_{(2)} \simeq \boldsymbol{\theta} - \frac{D_{ds}}{D_s}2\nabla_{\perp}\hat{\psi}(D_d\boldsymbol{\theta}), \tag{1.85}$$

(1) slowly varying $\simeq \frac{\zeta_s-\zeta_d}{\zeta_s} = \frac{D_{ds}}{D_s}$,

(2) $(\zeta_d\boldsymbol{\theta},\zeta) = (D_d\boldsymbol{\theta},\zeta)$,

where

$$\hat{\psi}(D_d\boldsymbol{\theta}) = \int d\zeta U(D_d\boldsymbol{\theta},\zeta). \tag{1.86}$$

Thus

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{ds}}{D_s}\hat{\boldsymbol{\alpha}}, \quad \hat{\boldsymbol{\alpha}} = 2\nabla_{\perp}\hat{\psi} = \frac{2}{D_d}\frac{\partial\tilde{\psi}}{\partial\boldsymbol{\theta}}, \tag{1.87}$$

with

$$\tilde{\psi}(\boldsymbol{\theta}) = \hat{\psi}(D_d\boldsymbol{\theta}). \tag{1.88}$$

From here, we again obtain

$$\nabla_\perp \cdot \hat{\boldsymbol{\alpha}} = 2\int \Delta_\perp U d\zeta = 8\pi G\Sigma, \tag{1.89}$$

$$\Delta_\perp \hat{\psi} = 4\pi G\Sigma. \tag{1.90}$$

Chapter 2

Simple lens models

It is now high time to study some simple, but important examples of specific types of lenses. Although they are simple, they turn out to be very useful to better understand the lensing phenomenon.

2.1 Axially symmetric lenses: generalities

If the lens is axially symmetric, our general lensing equations simplify considerably. For the deflection angle, this was already shown in Section 1.2. According to (1.36) we have then

$$\hat{\alpha}(\xi) = \frac{4G}{c^2}\frac{M(\xi)}{\xi} \tag{2.1}$$

(only the modulus of the angle counts). For the rescaled angle $\alpha(x)$ in (1.44) this translates to

$$\alpha(x) = \frac{m(x)}{x}, \qquad (\mathbf{x} = (x, 0), x > 0), \tag{2.2}$$

where

$$m(x) = 2\int_0^x \kappa(x')\, x' dx'. \tag{2.3}$$

The lens equation (1.46) can be written in scalar form

$$y = x - \alpha(x) = x - \frac{m(x)}{x}, \tag{2.4}$$

where now $x \in \mathbb{R}$ and $m(x) = m(|x|)$. From (1.48) we obtain

$$\alpha = \frac{d\psi}{dx}. \tag{2.5}$$

The Poisson equation (1.49) for ψ becomes

$$\frac{1}{x}\frac{d}{dx}\left(x\frac{d\psi}{dx}\right) = 2\kappa. \tag{2.6}$$

Inserting here (2.5) and (2.2) leads to

$$\frac{dm}{dx} = 2x\kappa(x), \tag{2.7}$$

which of course, follows also from (2.3). From (2.5), (2.2) and (2.3) we have

$$\frac{d\psi}{dx} = \frac{2}{x}\int_0^x \kappa(x')\, x' dx'. \tag{2.8}$$

In this equation, the right hand side is also equal to

$$\frac{d}{dx}2\int_0^x \kappa(x')\,x'\ln\left(\frac{x}{x'}\right)dx'.$$

Thus, provided that $\kappa(x)$decreases faster than x^{-1}, we find

$$\psi(x) = 2\int_0^x \kappa(x')\,x'\ln\left(\frac{x}{x'}\right)dx'. \tag{2.9}$$

Let us also look at the differential $D\varphi$ of the lens map. According to (2.4), φ is given by

$$\mathbf{y} = \mathbf{x} - \frac{m(x)}{x^2}\mathbf{x} \qquad (x = |\mathbf{x}|). \tag{2.10}$$

Hence,

$$D\varphi = 1 - \frac{m(x)}{x^4}\begin{pmatrix} x_2^2 - x_1^2 & -2x_1x_2 \\ -2x_1x_2 & x_1^2 - x_2^2 \end{pmatrix} - \frac{m'(x)}{x^3}\begin{pmatrix} x_1^2 & x_1x_2 \\ x_1x_2 & x_2^2 \end{pmatrix}. \tag{2.11}$$

Because of (2.7), the trace is correct (see (1.52)), and the components of the shear are

$$\gamma_1 = \frac{1}{2}\left(x_2^2 - x_1^2\right)\left(\frac{2m}{x^4} - \frac{m'}{x^3}\right), \qquad \gamma_2 = x_1x_2\left(\frac{m'}{x^3} - \frac{2m}{x^4}\right). \tag{2.12}$$

This gives

$$|\gamma|^2 = \left(\frac{m}{x^2} - \kappa\right)^2$$

and

$$\begin{aligned} \det D\varphi = & \ (1-\kappa)^2 - |\gamma|^2 = \\ & \left(1 - \frac{m}{x^2}\right)\left(1 + \frac{m}{x^2} - 2\kappa\right) = \\ & = \left(1 - \frac{1}{x}\frac{d\psi}{dx}\right)\left(1 - \frac{d^2\psi}{dx^2}\right). \end{aligned} \tag{2.13}$$

The last two factors are the eigenvalues of $D\varphi$.

This implies that there are two types of critical curves

$$\begin{aligned} \frac{m(x)}{x^2} &= 1: \quad \textit{tangential critical curve;} \\ \frac{d}{dx}\left(\frac{m(x)}{x}\right) &= 1: \quad \textit{radial critical curve.} \end{aligned} \tag{2.14}$$

(The terminology will soon become clear.) The image of a tangential critical curve degenerates according to (2.4) into the point $y = 0$ in the source plane.

We can look at the critical points on the x_1–axis with $\mathbf{x} = (x, 0)$, $x > 0$. Then

$$D\varphi = 1 - \frac{m(x)}{x^2}\begin{pmatrix} -1 & 0 \\ 0 & +1 \end{pmatrix} - \frac{m'}{x}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{2.15}$$

and this matrix must have an eigenvector $\mathbf{X}$ with eigenvalue zero. For symmetry reasons, the vector must be either tangential, $\mathbf{X} = (0,1)$, or normal, $\mathbf{X} = (1,0)$, to the critical curve (which must be a circle).

We see readily that the first case occurs for a tangential critical curve, and the second for a radial critical curve. It turns out that the radial critical curve consist of folds.

For a tangential critical curve $\{|\mathbf{x}| = x_t\}$, we have by (2.14) and (2.3)

$$m(x_t) = \int_0^{x_t} 2x\kappa(x)\,dx = x_t^2. \tag{2.16}$$

With (1.44) this translates to

$$\int_0^{\xi_t} 2\xi\Sigma\left(\xi\right)d\xi = \xi_t^2\Sigma_{crit}. \tag{2.17}$$

The total mass $M\left(\xi_t\right)$ inside the critical curve is thus

$$M\left(\xi_t\right) = \pi\xi_t^2\Sigma_{crit}. \tag{2.18}$$

This shows that the average density $\langle\Sigma\rangle_t$ inside the tangential critical curve is equal to the critical density,

$$\langle\Sigma\rangle_t = \Sigma_{crit}. \tag{2.19}$$

(Correspondingly, $\langle\kappa\rangle_t = 1$.)

This can be used to estimate the mass of a deflector if the lens is sufficiently strong and the geometry is such that almost complete Einstein rings are formed. If θ_{arc} is the angular distance of the arc, we obtain numerically

$$\begin{aligned} M\left(\langle\theta_{arc}\rangle\right) &= \pi\left(D_d\theta_{arc}\right)^2\Sigma_{crit} \approx \\ &\approx \left(1.1\times 10^{14}M_\odot\right)\left(\frac{\theta_{arc}}{30''}\right)^2\left(\frac{D}{1\text{ Gpc}}\right). \end{aligned} \tag{2.20}$$

(D is defined in (2.31) below.)

2.2 The Schwarzschild lens: microlensing

This is the simplest case and is most relevant for the MACHO search.

We shall soon see that a convenient length scale ξ_0 is provided by the Einstein radius

$$R_E = \sqrt{\frac{4GM}{c^2}\frac{D_dD_{ds}}{D_s}} = 610\ R_\odot\left[\frac{M}{M_\odot}\frac{D_s}{\text{kpc}}\frac{D_d}{D_s}\left(1-\frac{D_d}{D_s}\right)\right]^{\frac{1}{2}}. \tag{2.21}$$

Since $\Sigma\left(\boldsymbol{\xi}\right) = M\delta^2\left(\boldsymbol{\xi}\right)$, we then have $\kappa\left(\mathbf{x}\right) = \pi\delta^2\left(\mathbf{x}\right)$, according to (1.44), and thus $m\left(x\right)\equiv 1$. The latter equation follows also from (2.1), (2.2) and (1.44). Thus, since Eq. (2.5) implies $\psi\left(x\right) = \ln x$, we have

$$\alpha\left(x\right) = \frac{1}{x}, \tag{2.22}$$

and the lens map is given by

$$y = x - \frac{1}{x}. \tag{2.23}$$

If the source is on the symmetry axis ($y = 0$), then $x = \pm 1$ (Einstein ring). For a given source position y, (2.23) has two solutions

$$x_{1,2} = \frac{1}{2}\left(y \pm \sqrt{y^2+4}\right). \tag{2.24}$$

The magnification $\mu = \left|\det D\varphi\right|^{-1}$ follows immediately from (2.13)

$$\mu^{-1} = \left|1 - \frac{1}{x^4}\right|, \tag{2.25}$$

which gives for the two images

$$\mu_{1,2} = \frac{1}{4}\left|\frac{y}{\sqrt{y^2+4}} + \frac{\sqrt{y^2+4}}{y} \pm 2\right|. \tag{2.26}$$

The total magnification $\mu_p = \mu_1 + \mu_2$ is found to be

$$\mu_p = \frac{y^2 + 2}{y\sqrt{y^2 + 4}}. \tag{2.27}$$

It is this function which one observes for MACHOs. Note that

$$y = \eta/\eta_0 = \frac{\eta}{(D_s/D_d)\,\xi_0} = \frac{\eta/D_s}{\xi_0/D_d} = \beta\theta_E^{-1}, \tag{2.28}$$

where $\theta_E = R_E/D_d$, *i.e.*,

$$\theta_E = \left[\frac{4GM}{c^2}\frac{D_{ds}}{D_d D_s}\right]^{\frac{1}{2}} \tag{2.29}$$

is the angular separation corresponding to the Einstein radius, and β (see Fig. 1.2) is the angular separation of the source from the optical axis. Numerically, (2.29) reads

$$\begin{aligned} \theta_E &= (0.9'' \cdot 10^{-3}) \left(\frac{M}{M_\odot}\right)^{\frac{1}{2}} \left(\frac{D}{10\ \text{kpc}}\right)^{-\frac{1}{2}}, \\ &= (0.9'') \left(\frac{M}{10^{12} M_\odot}\right)^{\frac{1}{2}} \left(\frac{D}{1\ \text{Gpc}}\right)^{-\frac{1}{2}}, \end{aligned} \tag{2.30}$$

where

$$D := \frac{D_d D_s}{D_{ds}} \tag{2.31}$$

is the *effective distance*.

Even when it is not possible to see multiple images, the magnification can still be detected if the lens and source move relative to each other giving rise to lensing-induced time variability. This kind of variability is called *microlensing*. Microlensing was first observed in the multiple-imaged QSO2237+0305 [8]. As is well-known, Paczyński proposed in 1986 to monitor millions of stars in the LMC to look for such magnifications in a fraction of the sources. In the meantime, this has been successfully implemented. The time scale for microlensing-induced variations is obviously given by $t_0 = D_d\theta_E/v$, where v is a typical virial velocity of the galactic halo. Numerically

$$t_0 = 0.214\ \text{yr} \left(\frac{M}{M_\odot}\right)^{\frac{1}{2}} \left(\frac{D_d}{10\ \text{kpc}}\right)^{\frac{1}{2}} \left(\frac{D_{ds}}{D_s}\right)^{\frac{1}{2}} \left(\frac{200\ \text{km s}^{-1}}{v}\right). \tag{2.32}$$

(The ratio D_{ds}/D_s is close to unity.)

Typical light variation curves corresponding to (2.27) are shown in Fig. 2.1 and Fig. 2.2. Note that t_0 does not directly give the mass. The chance of seeing a microlensing event can be expressed in terms of the optical depth, defined as the probability that at any instant of time a given star is within the angle θ_E of a lens. This probability τ is given by

$$\tau = \frac{1}{d\Omega}\int dV n\,(D_d)\,\pi\theta_E^2, \tag{2.33}$$

where $dV = d\Omega D_d^2 dD_d$ is the volume of an infinitesimal spherical shell with radius D_d which covers a solid angle $d\Omega$. Indeed, the integral gives the solid angle covered by the Einstein circles of the lenses, and the probability τ is obtained upon dividing this quantity by the observed solid angle $d\Omega$. Inserting the expression (2.29) for θ_E gives

$$\tau = \int_0^{D_s} \frac{4\pi G\rho}{c^2}\frac{D_d D_{ds}}{D_s} dD_d = \frac{4\pi G}{c^2} D_s^2 \int_0^1 \rho\,(x)\,x\,(1-x)\,dx,$$

where $x = D_d D_s^{-1}$ and ρ is the mass density of the MACHOs. It is this density that determines τ.

Observations have shown that τ is a few times 10^{-6} (the statistical errors are still large). We will hear more about the present status of this exciting subject. The future looks promising.

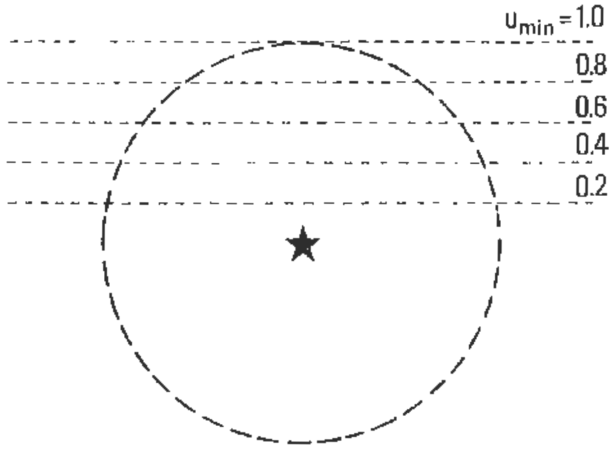

Figure 2.1: Einstein ring (dashed) and five possible relative orbits of a background star with projected minimal distances $u_{\min} = \xi_{\min}/R_E = 0.2, 0.4, ..., 1.0$.

2.3 Singular isothermal sphere

The so-called singular isothermal sphere is often used as a simple model for the mass distribution in elliptical galaxies. One arrives at this model by assuming an ideal isothermal gas law $p = (\rho/m)\, k_B T$ for the equation of state, where ρ is the mass density of stars and m the (average) mass of a star.

The equation of hydrostatic equilibrium then gives

$$\frac{k_B T}{m}\frac{d\rho}{dr} = -\rho\frac{GM(r)}{r^2}. \tag{2.34}$$

If we multiply this by $r^2\,(m/k_B T)$ and then differentiate with respect to r, we obtain, using

$$\frac{dM(r)}{dr} = 4\pi r^2 \rho, \tag{2.35}$$

the differential equation

$$\frac{d}{dr}\left(r^2\frac{d}{dr}\ln\rho\right) = -\frac{Gm}{k_B T}4\pi r^2\rho. \tag{2.36}$$

One arrives at this equation also in the kinetic theory as follows. Start from the Jeans equation, which one obtains by taking the first moment of the collisionless Boltzmann equation. In the stationary, spherically symmetric case this reads [9]

$$\frac{d}{dr}\left(n\sigma_r^2\right) + \frac{2n}{r}\left(\sigma_r^2 - \sigma_t^2\right) = -n\frac{dU}{dr}, \tag{2.37}$$

where n is the density of particles, and σ_r, σ_t are, respectively, the radial and transversal, velocity dispersions.

For the special case $\sigma_r^2 = \sigma_t^2 \equiv \sigma^2$ =const, Eq. (2.37) reduces to

$$\sigma^2\frac{dn}{dr} = -n\frac{GM\,(r)}{r^2}, \tag{2.38}$$

which is identical to (2.34) for $\rho = nm$, if we make the identification

$$k_B T = m\sigma^2. \tag{2.39}$$

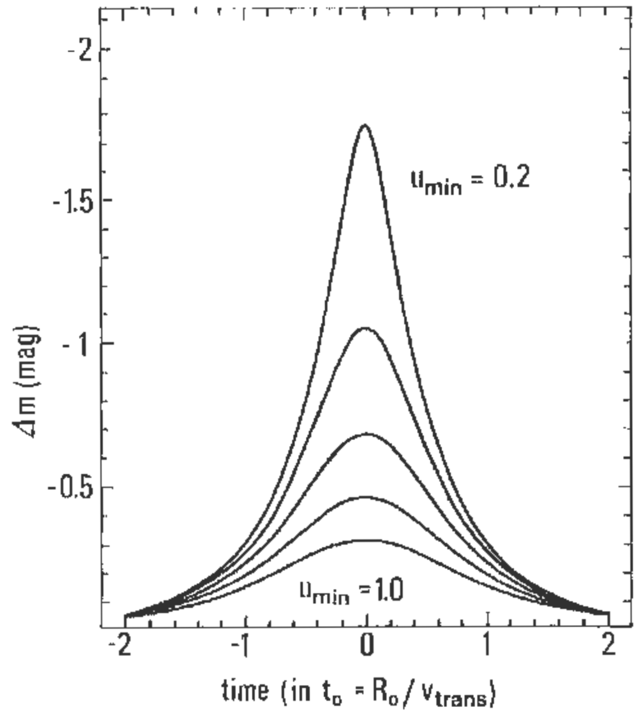

Figure 2.2: Light curves for the five cases in Fig. 2.1. The maximal magnification is $\mu = 1.34$ or $\Delta m = -0.32$ mag, if the star just touches the Einstein radius ($u_{\min} = 1.0$). For smaller values of $u_{\min}$ the maximal magnification is larger.

(Note, that σ^2 is the 1–dimensional velocity dispersion.) One solution, with a power dependence for $\rho(r)$, is easily found

$$\rho(r) = \frac{\sigma^2}{2\pi G r^2}. \tag{2.40}$$

Because the density is singular at the origin, it is called the *singular isothermal sphere* (regular solutions are only known numerically; [9]).

The projected mass density for (2.40) is easily found to be

$$\Sigma(\xi) = \frac{\sigma^2}{2G}\frac{1}{\xi}. \tag{2.41}$$

For the length scale ξ_0 we choose

$$\xi_0 = 4\pi\left(\frac{\sigma}{c}\right)^2 \frac{D_d D_{ds}}{D_s} \tag{2.42}$$

and obtain from (2.41) and (1.44)

$$\kappa(x) = \frac{1}{2|x|}, \tag{2.43}$$

thus

$$m(x) = |x|, \qquad \alpha(x) = \frac{x}{|x|} \quad (\Rightarrow \hat{\alpha} = 4\pi\sigma^2/c^2 = const), \tag{2.44}$$

and

$$y = x - \frac{x}{|x|}. \tag{2.45}$$

The Einstein ring is given by $|x| = 1, \quad |\xi| = \xi_0$. The corresponding angle is thus

$$\theta_E = 4\pi \left(\frac{\sigma}{c}\right)^2 \frac{D_{ds}}{D_s} \simeq (29'') \left(\frac{\sigma}{10^3 \text{ km s}^{-1}}\right)^2 \frac{D_{ds}}{D_s}. \tag{2.46}$$

This is often used also for clusters of galaxies.

We also note the mass $M\,(<\xi_0)$ inside the Einstein radius

$$\begin{aligned} M\,(<\xi_0) &= \int_0^1 \kappa\,(x)\,\Sigma_{crit} 2\pi x dx = \pi \xi_0^2 \Sigma_{crit} = \\ &= \pi\,(D_d \theta_E)^2\,\Sigma_{crit} \simeq 1.1 \times 10^{14} M_\odot \left(\frac{\theta_E}{30''}\right)^2 \left(\frac{D}{1 \text{ Gpc}}\right), \end{aligned} \tag{2.47}$$

where D is again given by (2.31).

The magnification for an image at x is easily found from (2.13)

$$\mu = \left|\frac{|x|}{|x| - 1}\right|. \tag{2.48}$$

For $|y| > 1$ one sees from (2.45) that there is only one image at $x = y + 1$ (take $y > 0$). When $|y| < 1$, there are two images, at $x = y + 1$ and $x = y - 1$. Using this in (2.48) we find for the total magnification of a point source

$$\mu_p = \begin{cases} 2/y & for \quad y \leq 1, \\ (1+y)/y & for \quad y \geq 1. \end{cases} \tag{2.49}$$

Note that the inner image becomes very faint for $y \to 1$.

Finally, we determine the time delay for the two images. This is easily found from (1.61) and $\psi\,(x) = |x|$ (from (2.8))

$$c\Delta t = \left[4\pi \left(\frac{\sigma}{c}\right)^2\right]^2 \frac{D_d D_{ds}}{D_s}\, 2y. \tag{2.50}$$

(Recall that no red-shift effects have been taken into account.)

2.4 Isothermal sphere with finite core radius

Since no analytic solutions of (2.36) without a central singularity are known, the surface mass density (2.41) is often modified parametrically by introducing a finite core radius ξ_c

$$\Sigma\,(\xi) = \frac{\sigma^2}{2G} \frac{1}{\sqrt{\xi^2 + \xi_c^2}}. \tag{2.51}$$

Using the same scale length ξ_0 as before, the corresponding dimensionless surface mass density is

$$\kappa\,(x) = \frac{1}{2\sqrt{x^2 + x_c^2}}, \qquad x_c := \xi_c/\xi_0. \tag{2.52}$$

With the help of the formulas in Section 2.1, you can easily work out every thing you like. When does the lens become critical? Consider, in particular, an *extended* source near $y = 0$ and discuss the form of the images (arcs and counter arcs, etc.).

In Chapter 8 of SEF other examples are worked out.

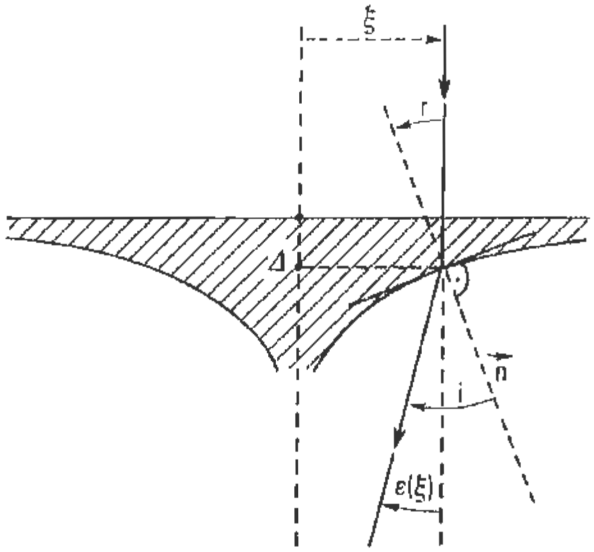

Figure 2.3: Deflection of a light ray passing through an axially symmetric lens (taken from S. Refsdal and J. Surdej, 1994 [10]).

2.5 Lensing experiments

I show now how one can mimic the deflection of light rays by an axially symmetric gravitational lens with an optical lens. Hopefully, some of you will construct such a lens for didactical purposes.

For simplicity, we choose the lens such that it is flat on one side and determine the axially symmetric shape of the other side such that the deflection angle comes out right. All angles in Fig. 2.3 can be assumed to be small. If n denotes the refraction index of the lens with respect to the air, we have

$$\frac{\sin(i)}{\sin(r)} = n \approx \frac{i}{r}$$

The deflection angle $\varepsilon(\xi) = i - r \simeq r(n-1)$ should agree with $\hat{\alpha}(\xi) = 4GM(\xi)/c^2\xi$:

$$r(n-1) = \frac{4GM(\xi)}{c^2\xi}.$$

From Fig. 2.3 we read off that

$$\frac{d\Delta}{d\xi} = -tg(r) \approx -r.$$

The shape of the optical lens is thus determined by

$$\frac{d\Delta}{d\xi} = -\frac{4G}{c^2(n-1)}\frac{M(\xi)}{\xi}, \qquad \Delta(\infty) = 0. \tag{2.53}$$

Let us first mimic the Schwarzschild lens, $M(\xi) = M$. In this case we get from (2.53)

$$\Delta(\xi) = -\frac{4GM}{c^2(n-1)}\ln\xi + const = \Delta(\xi_0) + \frac{2R_s}{n-1}\ln\left(\frac{\xi_0}{\xi}\right), \tag{2.54}$$

where R_s is the Schwarzschild radius. The choice of $(\xi_0, \Delta(\xi_0))$ is up to you. Example: plexiglass with $n = 1.49$, $\xi_0 = 14$ cm, $\Delta(\xi_0) = 1$ cm, $R_s = 0.3$ cm (corresponding to one third of the Earth mass). The shape is shown in Fig. 2.4a. Let us determine also the shape corresponding to the

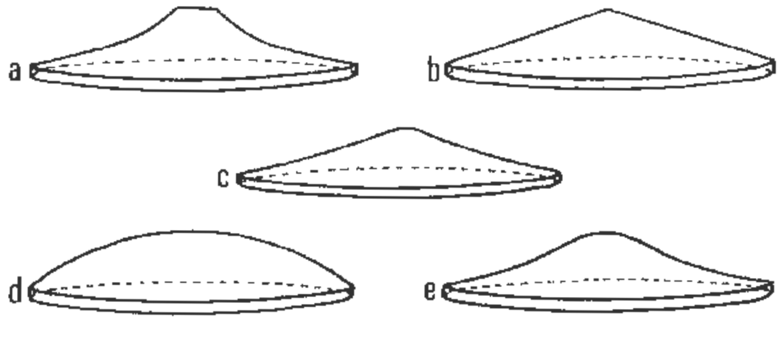

Figure 2.4: Several examples of axially symmetric optical lenses simulating the light deflection properties due to: (a) a point mass, (b) a singular isothermal sphere, (c) a spiral galaxy, (d) a uniform disk, (e) a truncated uniform disk of matter (taken from S. Refsdal and J. Surdej, Rep. Progr. Phys. 1994 [10]).

singular isothermal lens. From (2.41) we obtain

$$M(\xi) = \frac{\pi\sigma^2}{G}\xi \equiv K\xi, \tag{2.55}$$

hence

$$\Delta(\xi) = \Delta(\xi_0) + K(\xi_0 - \xi). \tag{2.56}$$

This is an axially symmetric cone (Fig. 2.4b).

Other examples are carried out in the review article by S. Refsdal and J. Surdej [10]. These authors describe also a convenient experimental set up.

2.6 Extended source

We go back to a general axially symmetric lens (Section 2.1), and study the magnification of a source close to the symmetry axis, assuming that there exists a tangential critical curve at $x_t\ (m(x_t) = x_t^2)$. Let us first consider a point source near $y = 0$. The lens equation (2.4),

$$y = x - \frac{m(x)}{x},$$

has two solutions close to the critical curve at $x = x_t + \Delta x$, $\quad x = -x_t + \Delta x$, where $\Delta x = \left(\frac{dy}{dx}\right)^{-1} y$.

Now, we have $det\ D\varphi = \frac{y}{x}\frac{dy}{dx}, \frac{y}{x} = 1 - \bar{\kappa}(x)$ with $\bar{\kappa}(x) \equiv \frac{m(x)}{x^2}$ (= mean mass surface density within x). Thus

$$|\det D\varphi| \simeq |1 - \bar{\kappa}(\pm x_t + \Delta x)|\,|dy/dx| \simeq \left|\frac{d\bar{\kappa}}{dx}\right|_{x_t} |y|$$

and so the total magnification of the point source is

$$\mu_p = g_p/y, \qquad g_p = 2\left|\frac{d\bar{\kappa}}{dx}(x_t)\right|^{-1}. \tag{2.57}$$

Now, we turn to an extended source with surface brightness profile $I(\mathbf{y})$. Its magnification μ_e is generally given by

$$\mu_e = \frac{\int I(\mathbf{y})\,\mu_p(\mathbf{y})\,d^2y}{\int I(y)\,d^2y}. \tag{2.58}$$

For an axially symmetric lens in polar coordinates centered on a circular source with radius R and brightness distribution $I(r/R)$, we have

$$\mu_e = \left[2\pi \int_0^\infty I(r/R)\, r dr\right]^{-1} \int_0^\infty r dr I(r/R) \int_0^{2\pi} d\varphi \frac{g_p}{\sqrt{y^2+r^2+2ry\cos\varphi}}.$$

Thus

$$\mu_e = \frac{g_p}{R}\zeta(y/R) \tag{2.59}$$

where

$$\zeta(u) = \left[\pi \int_0^\infty I(x)\, x dx\right]^{-1} \cdot \int_0^\infty I(x)\, x dx \int_0^\pi \frac{d\varphi}{\sqrt{u^2+x^2+2ux\cos\varphi}}. \tag{2.60}$$

The function $\zeta(u)$ is discussed in SEF, p.238. For a uniform brightness its maximum is at $u=0$ where $\zeta(0)=2$. At any rate, μ_e remains finite for $y \to 0$. If one computes, as an exercise, $\mu_e^{\max}$ for Schwarzschild lens, one finds, using (2.27),

$$\mu_e^{\max} = \frac{\sqrt{4+R^2}}{R}. \tag{2.61}$$

2.7 Two point–mass lens

This is an instructive non–axially symmetric lens model. It has also become highly relevant recently, because binary microlensing events (OGLE #7, DUO #2,......) have been discovered (see Fig. 2.5). For several point masses M_i at transversal positions $\boldsymbol{\xi}_i$ the general formula (1.34) for the

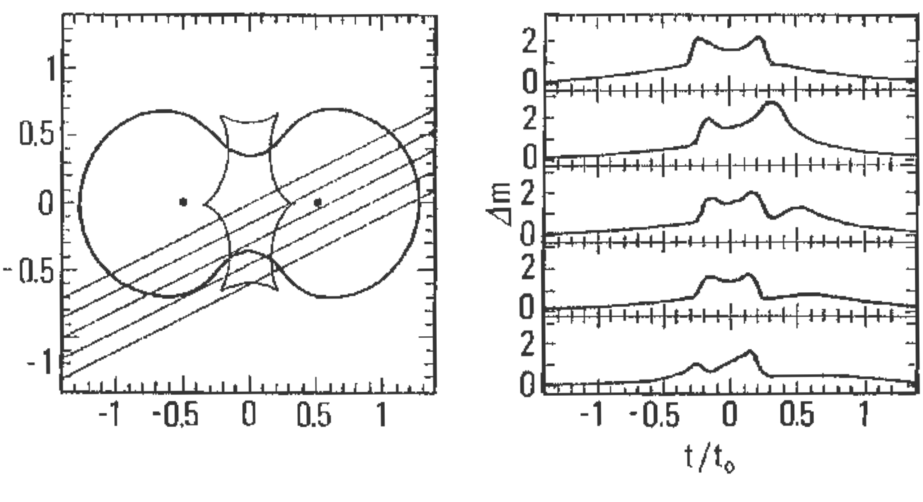

Figure 2.5: Light curve of a binary microlensig event (taken from R. Narayan and M. Bartelmann, Lectures on Gravitational Lensing, 1995 [11]).

deflection angle gives

$$\hat{\alpha}(\boldsymbol{\xi}) = \sum_{i=1}^{N} \frac{4GM_i}{c^2} \frac{\boldsymbol{\xi}-\boldsymbol{\xi}_i}{|\boldsymbol{\xi}-\boldsymbol{\xi}_i|^2}. \tag{2.62}$$

Let $M = \sum\limits_{i=1}^{N} M_i$ be the total mass and $M_i = \mu_i M$. For the length scale ξ_0 we choose the Einstein

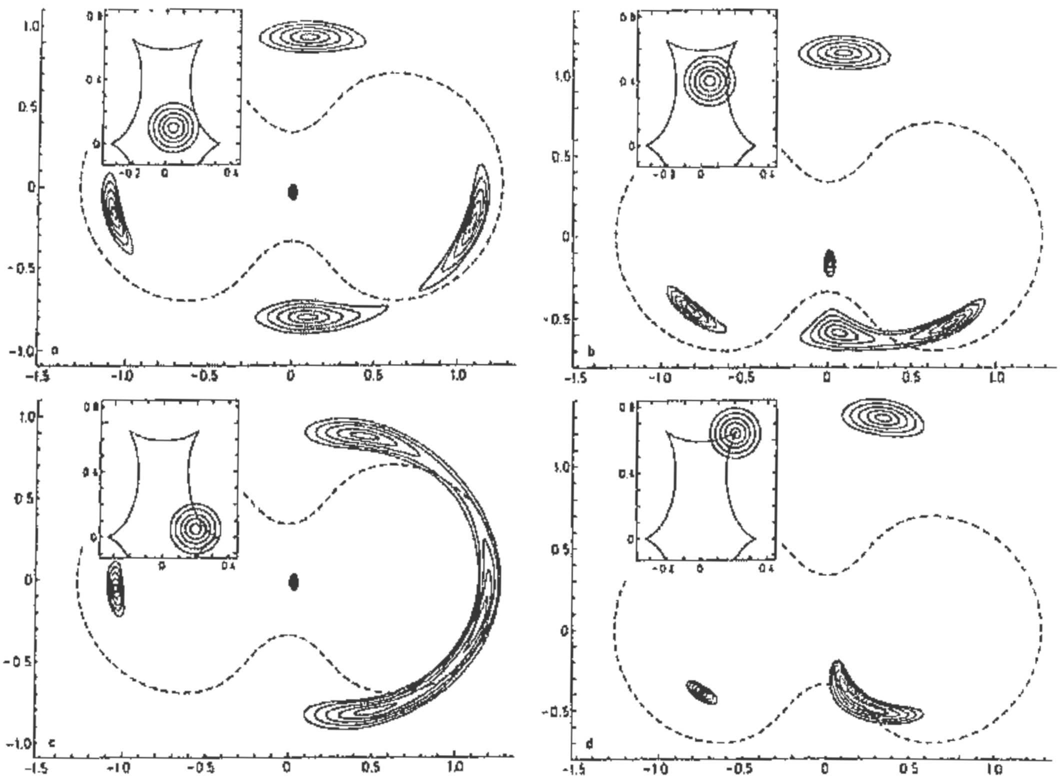

Figure 2.6: Imaging of extended sources for the two point–mass lens with lens separation $2\chi = 1.0$. The inserts show the isophotes of a circular source, together with part of the caustic. The dashed line is the critical curve. Depending on the source position, the images can have vastly different shapes (taken from SEF, p. 263).

radius (2.21) for the total mass. Then the lens map becomes

$$\mathbf{y} = \mathbf{x} - \sum_{i=1}^{N} \frac{\mu_i}{\left|\mathbf{x} - \mathbf{x}_i\right|^2} \left(\mathbf{x} - \mathbf{x}_i\right), \tag{2.63}$$

where $\mathbf{x}_i = \boldsymbol{\xi}_i / \xi_0$.

Even for the two point–mass lens the analysis of this map is fairly complicated. Fig. 2.6, taken from SEF (p. 263), is very instructive. Depending on the position of the source relative to the caustic, the image shape varies strongly. If the source lies completely inside the caustic, five separate images are produced (one is very close to the line connecting the two masses). The two images close to the critical curve are highly elongated and point towards each other. In (b) the source lies on a caustic and this leads to the formation of images with internal structure, because the inner isophotes still have five separate images, while those which cross the caustic have fewer. More complicated images are formed (c) close to a cusp. You should have a careful look at the figure.

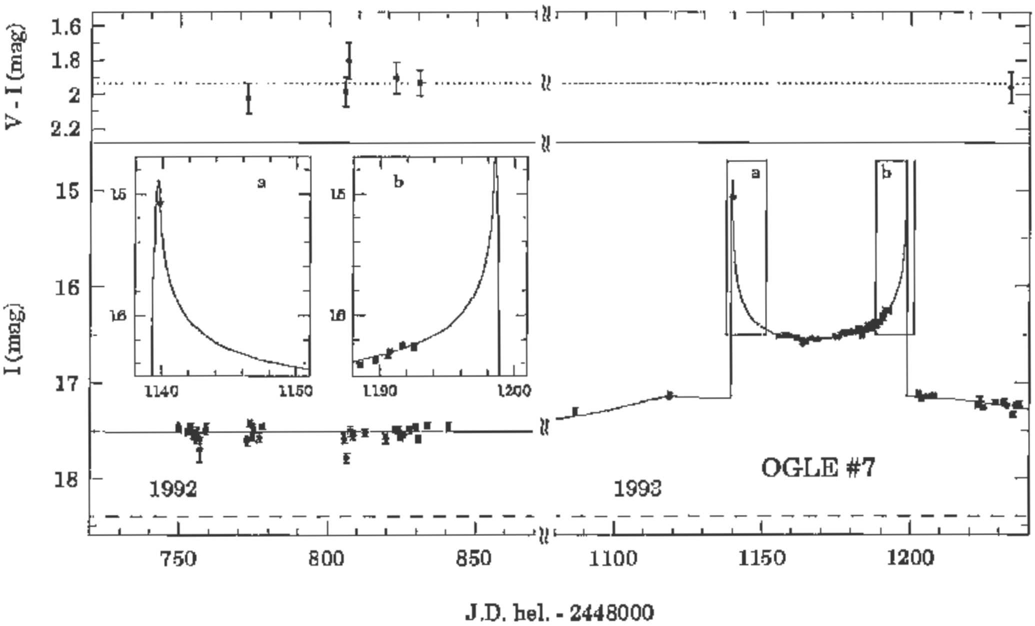

Figure 2.7: Light curve of the first binary microlensing event, OGLE # 7 (taken from the OGLE web page http://www.astro.princeton.edu/ stanek/ogle/).

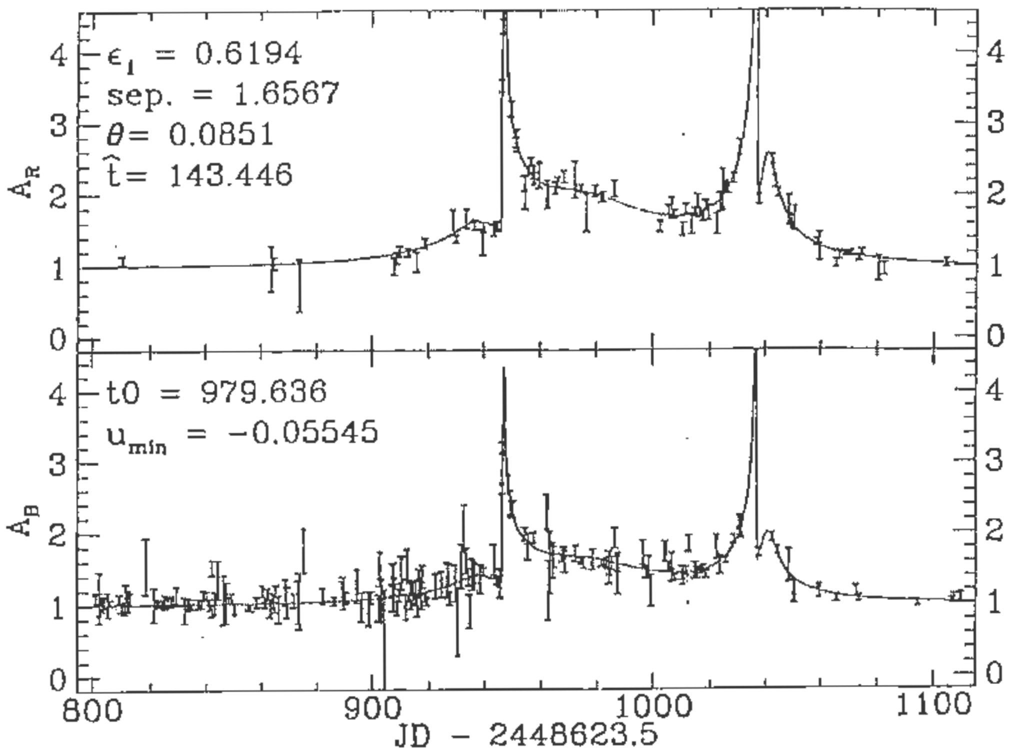

Figure 2.8: Binary microlensing event (taken from MACHO Collaboration web page http://darkstar.astro.washington.edu).

Chapter 3

Lensing by galaxy clusters

Clusters of galaxies show two types of lensing phenomena (see Fig. 1):

(i) Rich centrally condensed clusters produce sometimes giant arcs when a background galaxy happens to be almost aligned with one of the cluster caustics (strong lensing). A famous case is the circular shape arc in Cl2244-02, shown in Fig. 3.1, which extends over more than 100°.

(ii) Every cluster produces weakly distorted images of a large number of background galaxies (weak lensing). A beautiful example is Abell 2218 (Fig. 1).

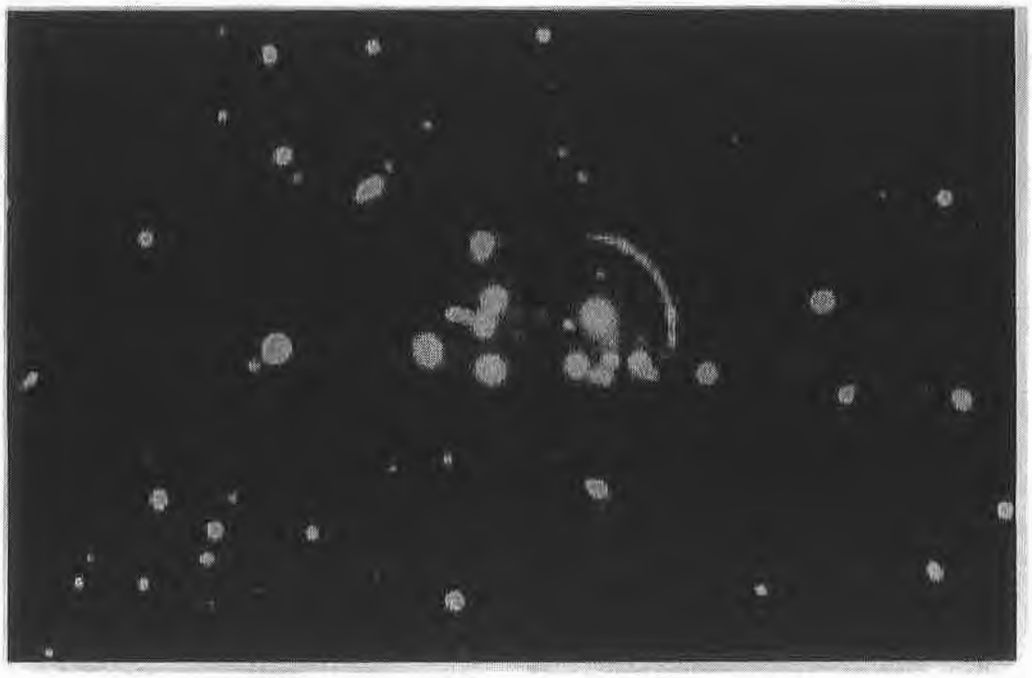

Figure 3.1: Giant arc in Cl2244-02 (ultra–deep B image from CFHT). The lensing cluster is at $z = 0.329$ and the source of the arc is a very distant field galaxy at $z = 2.238$. The stellar content seems normal. (Courtesy of G. Soucail, Obs. Midi–Pyrénées, ESO Messenger 69, September 1992.)

For the analysis of giant arcs, we have to use (unfortunately) *parametrized* lens models which are fitted to the observational data. The situation is much better for weak lensing, because there now exist several *parameter–free reconstruction* methods of projected mass distributions from weak lensing data. (For a review, see [29].) In this chapter I will concentrate on this topical issue, and provide some of the theoretical background.

3.1 Strong lensing by clusters

Strong lensing occurs when in the central region $\Sigma(\boldsymbol{\xi})$ becomes larger than the critical surface mass density (1.45). In the ideal case of an axisymmetric mass distribution and a source right behind the lens' center, the deflection angle becomes

$$\hat{\alpha}(\theta) = \frac{4G}{c^2} \frac{M(<\theta)}{D_d \theta} \tag{3.1}$$

(see (1.36)), and the lens equation (1.43) reduces to

$$\theta D_s = \hat{\alpha}(\theta) D_{ds}. \tag{3.2}$$

These two equations give

$$M(<\theta) = \pi (\theta D_d)^2 \Sigma_{crit}. \tag{3.3}$$

The radius θ_{arc} of a large arc gives an estimate of the Einstein radius of the cluster and (3.3) for $\theta = \theta_{arc}$ provides an estimate of the mass enclosed by the arc, if we know the redshifts of the lens and the source. Even if no ring–shaped image is produced, a mass estimate with this procedure is useful and often surprisingly accurate. For example, a quadrupole image system, such as the "Einstein cross" (QSO2237+0305) allows one to trace approximately the Einstein "circle" and a mass estimate can be obtained with (3.3). This is, however, only the first step. For extended sources detailed modellings have been made making use of elaborate techniques. (See, e.g. [12], and references therein.)

One can also get a simple estimate of the velocity dispersion σ_v by using (2.49), *i.e.*,

$$\sigma_v \simeq 10^3 \text{ km s}^{-1} \left(\frac{\theta_{\mathrm{E}}}{29''}\right)^{\frac{1}{2}} \left(\frac{\mathrm{D_d}}{\mathrm{D_{ds}}}\right)^{\frac{1}{2}}. \tag{3.4}$$

Table 1 lists masses, mass–to–blue–light ratios, and velocity dispersions of three clusters with prominent arcs. Further results can be found in the review article by Fort and Mellier [13].

Let me emphasize the limitations of all this. First the analysis is *model dependent* and one determines the mass only inside a cylinder of the inner part of a lensing cluster.

As a historical footnote, I should add that the discovery of arcs was a surprise, because people thought that clusters are not compact enough to produce critical curves. This was based on estimates of core radii from X–ray observations of the intercluster gas which came out larger than required for critical cluster. This discrepancy between core radii $\leq 30\ h^{-1}$ kpc inferred from arcs and the results of X–ray imaging has been discussed a lot, and various explanations have been put forward (see, e.g., Ref. [14]).

Generally speaking, the dark matter and hot gas density profiles do not have to follow each other. In particular, an isothermal X–ray gas in hydrostatic equilibrium may develop a flat core well outside the radius where giant arcs form. Several reasons, like projection effects, have been suggested to explain the apparent discrepancies [15], [16].

Another interesting result is worth mentioning. Numerically generated cluster mass profiles by Bartelmann, Steinmetz and Weiss [17] show that the probability for forming arcs in these clusters is *substantially higher* than that of more symmetric mass profiles with the same mass. Asymmetries and substructure increase the total length of the caustic curves. This is probably related to the fact that the shear is increased by substructure, implying that critical curves can occur also in regions where κ is less than unity (see [18]).

A remarkable phenomenon is the occurrence of so–called *radial arcs* in galaxy clusters. These are *radially* rather than tangentially elongated, as most luminous arcs are. They are much less numerous (examples: MS 2137, Abell 370). Their position has been interpreted in terms of the turnover of the mass profile and a core radius $\sim 20\ h^{-1}$ kpc has been deduced, quite independent of any details of the lens model. There are, however, other mass profiles which can produce radial arcs, and have no flat core; even singular density profiles can explain radial arcs [19]. Such singular profiles of the dark matter are consistent with the large core radii inferred from X–ray emission (see [14]).

Table 1: Masses, mass–to–blue–light ratios, and velocity dispersions for three clusters with prominent arcs.

Cluster	M ($M_\odot$)	M/L_B (solar)	σ(km s^{-1})
A370	$\sim 10^{14}$	~ 200	~ 1550
A2390	$\sim 1.5 \times 10^{14}$	~ 120	~ 1250
MS2137-23	$\sim 5 \times 10^{13}$	~ 250	~ 1100

3.2 Mass reconstruction from weak lensing

There is a population of distant blue galaxies in the universe whose spatial density reaches 50–100 galaxies per square arc minute at faint magnitudes. The images of these distant galaxies are *coherently distorted* by any forground cluster of galaxies. Since they cover the sky so densely, the *distortions can be determined statistically* (individual weak distortions cannot be determined, since galaxies are not intrinsically round). Typical separations between arclets are $\sim (5 \div 10)''$ and this is much smaller than the scale over which the gravitational cluster potential changes appreciably.

Initiated by an influential paper of Kaiser and Squires [20], a considerable amount of theoretical work on various parameter–free reconstruction methods has recently been carried out [21] [22]. The main problem consists in the task to make optimal use of limited noisy data, without modeling the lens.

Parameter–free inversions can most simply be described by making use of a complex formulation of lensing theory. In this formalism, the relevant equations emerge almost automatically [23]. I will describe this in detail in Chapter 5. In what follows, I shall, however, only use what we have learned in the previous sections.

3.2.1 Relations between mean convergence and reduced shear

The reduced shear g, introduced in Section 1.7, is in principle observable over a large region. What we are really interested in, however, is the mean curvature κ, which is related to the surface mass density by (1.44). Since by (1.77)

$$\mathrm{g} = -\frac{\gamma}{1-\kappa} \tag{3.5}$$

we first look for relations between the shear γ and κ.

Recall that

$$\kappa = \frac{1}{2}\Delta\psi, \qquad \gamma_1 = \frac{1}{2}\left(\psi_{,11} - \psi_{,22}\right) \equiv D_1\psi, \qquad \gamma_2 = \psi_{,12} \equiv D_2\psi, \tag{3.6}$$

where

$$D_1 := \frac{1}{2}\left(\partial_1^2 - \partial_2^2\right), \qquad D_2 := \partial_1\partial_2. \tag{3.7}$$

Note the identity

$$D_1^2 + D_2^2 = \frac{1}{4}\Delta^2. \tag{3.8}$$

Hence

$$\Delta\kappa = 2\sum_{i=1,2} D_i\gamma_i. \tag{3.9}$$

Here, we can substitute the reduced shear, given by Eq. (3.5), on the right for γ. This gives the important equation

$$\Delta\kappa = -2\sum_i D_i[g_i\,(1-\kappa)]. \tag{3.10}$$

For a given (measured) $\mathbf{g}$ this equation does not determine uniquely κ, which is a famous *mass-sheet degeneracy* (a homogeneous mass sheet does not produce any shear). For a given $\mathbf{g}$, Eq. (3.10) remains invariant under the substitution

$$\kappa \to \lambda\kappa + (1-\lambda)\,, \tag{3.11}$$

where λ is a real constant.

Eq. (3.10) can be turned into an integral equation, by making use of the fundamental solution

$$\mathcal{G} = \frac{1}{2\pi}\ln|\mathbf{x}|\,, \tag{3.12}$$

introduced in (1.37). One solution of (3.9) is

$$\kappa = 2\mathcal{G} * \left(\sum_i D_i\gamma_i\right) + \kappa_0, \tag{3.13}$$

with a real constant κ_0. The most general solution corresponds to κ_0 replaced any harmonic function. For physical reasons, this function must, however, be bounded, and it is a constant. Replacing γ again by the reduced shear, we obtain an integral equation for κ. We write this in a different form by noting that

$$D_1 \ln|\mathbf{x}| = \frac{x_2^2 - x_1^2}{|\mathbf{x}|^4} \equiv \mathcal{D}_1, \qquad D_2 \ln|\mathbf{x}| = -\frac{2x_1x_2}{|\mathbf{x}|^4} \equiv \mathcal{D}_2. \tag{3.14}$$

Since

$$\mathcal{G} * (D_i\gamma_i) = (D_i\mathcal{G}) * \gamma_i = \frac{1}{2\pi}\mathcal{D}_i * \gamma_i,$$

we obtain from (3.13)

$$\kappa = \kappa_0 + \frac{1}{\pi}\left[\mathcal{D}_1 * \gamma_1 + \mathcal{D}_2 * \gamma_2\right], \tag{3.15}$$

and thus the integral equation

$$\kappa = \kappa_0 - \frac{1}{\pi}\mathcal{D}_1 * \left[g_1(1-\kappa)\right] - \frac{1}{\pi}\mathcal{D}_2 * \left[g_2(1-\kappa)\right]. \tag{3.16}$$

Eq. (3.15) appears the first time in [20]. The integral equation (3.16) has been used, for instance, in [22] for nonlinear cluster inversions.

Note also

$$\nabla\kappa = \begin{pmatrix} \kappa_{,1} \\ \kappa_{,2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(\psi_{,111} + \psi_{,221}) \\ \frac{1}{2}(\psi_{,112} + \psi_{,222}) \end{pmatrix} = \begin{pmatrix} \gamma_{1,1} + \gamma_{2,2} \\ \gamma_{2,1} - \gamma_{1,2} \end{pmatrix}. \tag{3.17}$$

This expression for the gradient $\nabla\kappa$ in terms of the shear has been translated by Kaiser [24] into a relation involving the reduced shear.

We proceed as follows. Let $K = \ln(1-\kappa)$, then

$$\nabla K = -(1-\kappa)^{-1}\nabla\kappa. \tag{3.18}$$

In addition, we have

$$\partial_j\gamma_i = -\partial_j\left[(1-\kappa)g_i\right] = -(1-\kappa)\partial_j g_i + g_i\partial_j\kappa,$$

and thus, with (3.17),

$$\partial_1\kappa = \partial_1\gamma_1 + \partial_2\gamma_2 = -(1-\kappa)(\partial_1 g_1 + \partial_2 g_2) + g_1\partial_1\kappa + g_2\partial_2\kappa.$$

Hence, we have

$$(1-g_1)\frac{\partial_1\kappa}{1-\kappa} - g_2\frac{\partial_2\kappa}{1-\kappa} = -\partial_1 g_1 - \partial_2 g_2.$$

Similarly,

$$-g_2\frac{\partial_1\kappa}{1-\kappa}+(1+g_1)\frac{\partial_2\kappa}{1-\kappa}=-\partial_1 g_2+\partial_2 g_1.$$

The left hand side of this linear system for $\nabla\kappa/(1-\kappa)$ is given by the matrix

$$M=\begin{pmatrix}1-g_1 & -g_2\\ -g_2 & 1+g_1\end{pmatrix}, \tag{3.19}$$

with the inverse

$$M^{-1}=\frac{1}{1-g_1^2-g_2^2}\begin{pmatrix}1+g_1 & g_2\\ g_2 & 1-g_1\end{pmatrix}. \tag{3.20}$$

This, together with (3.18), gives

$$\boldsymbol{\nabla} K=u, \tag{3.21}$$

where

$$u=\frac{1}{1-g_1^2-g_2^2}\begin{pmatrix}1+g_1 & g_2\\ g_2 & 1-g_1\end{pmatrix}\begin{pmatrix}\partial_1 g_1+\partial_2 g_2\\ \partial_1 g_2-\partial_2 g_1\end{pmatrix}. \tag{3.22}$$

In principle, the gradient of $K=\ln(1-\kappa)$ is thus observable.

As was emphasized earlier, this can be done only statistically, by determining $\langle \mathrm{g}\rangle=-<\gamma/(1-\kappa)>$ (for $\kappa<<1$, $\langle \mathrm{g}\rangle\simeq -<\gamma>$).

3.2.2 Practical difficulties, examples

In practice, there are several difficulties that complicate the application of the inversion formulas derived so far. I do not have to tell astronomers that atmospheric turbulence causes images to be blurred and thus circularize elliptical images taken by ground–based telescopes. On the other hand, anisotropies of the point–spread function can introduce spurious ellipticities. These effects have to be taken into account with high precision.

Another difficulty is that reconstruction equations of the type (3.15) and (3.16) involve convolutions over the *entire* plane. Real lensing data are, however, always confined to a *finite field* of the sky. For this reason, it is important to find integral formulas in which only integrations over bounded domains occur. This is possible [22], as I show in Chapter 5, by using the complex formulation.[1]

[1]In an ideal world (without measuring errors) equation (3.21),

$$\boldsymbol{\nabla} K=\mathbf{u}, \tag{3.23}$$

(with $\mathbf{u}$ given by (3.22) in terms of the reduced shear) can be solved in various ways. For instance, by taking the divergence, we get

$$\Delta K=\boldsymbol{\nabla}\cdot\mathbf{u}\quad \text{inside } \Omega, \tag{3.24}$$

and by taking the scalar product of (3.23) with the outward unit normal $\mathbf{n}$ on $\partial\Omega$, we find for the normal derivative

$$\frac{\partial K}{\partial u}=\mathbf{n}\cdot\mathbf{u}\quad \text{on } \partial\Omega. \tag{3.25}$$

Equations (3.24) and (3.25) constitute a Neumanu boundary problem for K, which determines K up to a constant. There are efficient and fast methods for a numerical solution of the Neumann problem.

In reality, however, the vector field $\mathbf{u}$ comes from uoisy observational data, and hence will *not* be a gradient field. We can, of course, consider a decomposition of $\mathbf{u}^{(\mathrm{obs.})}$ into a gradient and a rotational part,

$$u_i^{(\mathrm{obs.})}=\partial_i\tilde{K}+\varepsilon_{ij}\partial_j s, \tag{3.26}$$

but this is uot uuique on Ω. The question is which $\boldsymbol{\nabla}\tilde{K}$ should be identified with $\boldsymbol{\nabla} K$.

The rotational part in (3.26) is due to noise and we naturally impose the condition that its mean over Ω vanishes. (It should, of course, also vanish if u is already a gradient field.) By Gauss' theorem, this condition is equivalent to

$$\int_{\partial\Omega} s\varepsilon_{ij}n_j\,dl=0.$$

A sufficient condition for this is $s\,|_{\partial\Omega}=\text{const.}$ (nse againt Gauss' theorem). Then we have besides

$$\Delta\tilde{K}=\boldsymbol{\nabla}\cdot\mathbf{u}^{(\mathrm{obs.})}$$

Once the mean curvature $\kappa(\mathbf{x})$ has been reconstructed, we would need red shifts of the source galaxies. These are, however, *not available*. The induced uncertainty is not very serious for low redshift clusters, because the influence becomes weak if the sources are at much higher redshifts than the cluster. (Note that the relation between κ and $\sum$ involves the ratio D_s/D_{ds}.) For a detailed discussion of including the redshift distribution, see [26].

The mass–sheet degeneracy is broken also if we have information, for instance, on the magnification μ, since this transforms as $\mu \to \mu/\lambda^2$ under (3.11); use $\gamma \to \lambda\gamma$ and $\mu = \left[(1-\kappa)^2 - \gamma^2\right]^{-1}$. The magnification is accessible on a statistical basis by comparing the sizes of galaxies in cluster fields with those of galaxies of equal surface brightness in empty fields, or by the change in number density of galaxies.

We now give some results on cluster reconstructions from weak lensing data.

A typical example is shown in Fig. 3.2. On the left one sees a HST image of the cluster Cl0024, overlaid with the shear field obtained from an observation of arclets with the CFHT [27], and on the right the reconstructed surface–mass density from the shear field [28]. This reconstruction was obtained with the help of a non–linear, finite–field algorithm (see also the review [11]).

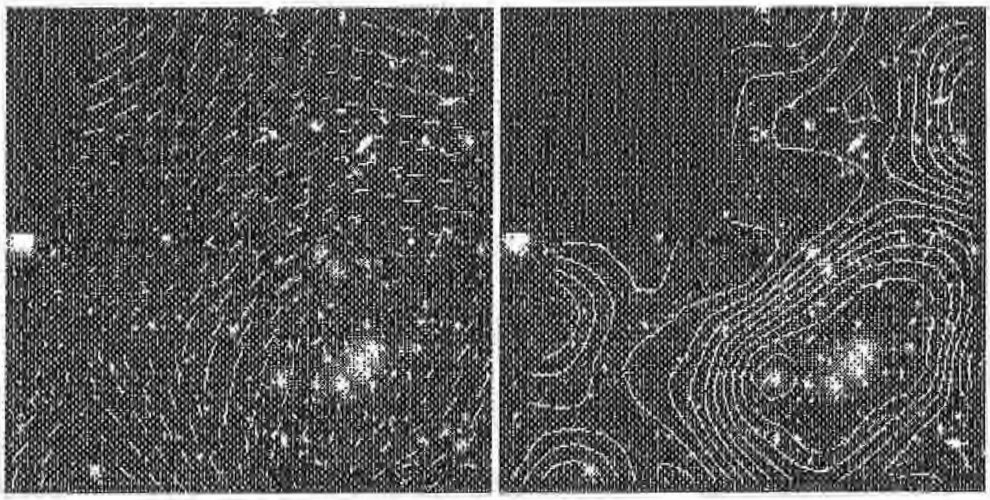

Figure 3.2: Shear field and surface-mass reconstruction of Cl0024 (taken from R. Narayan and M. Bartelmann, Lectures on Gravitational Lensing, 1995 [11]).

3.3 Comparison with results from X–ray observations

Beside the lensing technique, there are two other methods for determining mass distributions of clusters:
1) The observed velocity dispersion, combined with the Jeans–equation from stellar dynamics (Eq. (2.39)) gives the total mass distribution, if it is assumed that light traces mass.
2) X–ray observations of the intracluster gas, combined with the condition of hydrostatic equilibrium and spherical symmetry leads also to the total mass distribution as well as to the baryonic distribution.

also

$$\frac{\partial \tilde{K}}{\partial n} = \mathbf{n}\cdot\mathbf{u}^{(\mathrm{obs.})},$$

because $n_i\varepsilon_{ij}\partial_j s = 0$ (∇s is parallel to $\mathbf{n}$).

These considerations led Seitz and Schneider [25] to identify ∇K with the solution of the Neumann problem (3.24), (3.25) with the experimental (noisy) data $\mathbf{u}^{(\mathrm{obs.})}$. For tests of the method, see their paper.

This second method is very topical and will also play an important role in future developments. It may, therefore, be appropriate to remind you of the main points. If the hydrostatic equilibrium equation for the hot gas

$$\frac{dP_g}{dr} = -\rho_g \frac{GM_t(r)}{r^2}$$

is combined with the ideal equation of state $P_g = (k_B T_g/\mu m_H)\,\rho_g$, one easily finds for the total mass profile

$$M_t(r) = -\frac{k_B T_g}{G\mu m_H}\left(\frac{d\ln\rho_g}{d\ln r} + \frac{d\ln T_g}{d\ln r}\right) r. \tag{3.27}$$

The right hand side can be determined from the intensity distribution and some spectral information. (At present, the latter is not yet good enough, because of relatively poor resolution, which, however, will change with the XMM survey.) Note, that we assumed spherical symmetry, and this can introduce substantial errors if the cluster is strongly elongated along the line of sight.

Weak lensing, together with an analysis of X–ray observations offers a unique possibility to probe the *relative distributions* of the gas and the dark matter, and to study the dynamical relationship between the two.

Let me show an example of such a comparison. The cluster of galaxies A2163 (z=0.201) is the hottest and one of the two most massive known so far. GINGA satellite measurements gave an X–ray temperature of $\sim$ 14 keV and a X–ray luminosity of 6×10^{45} erg s^{-1}. ROSAT measurements

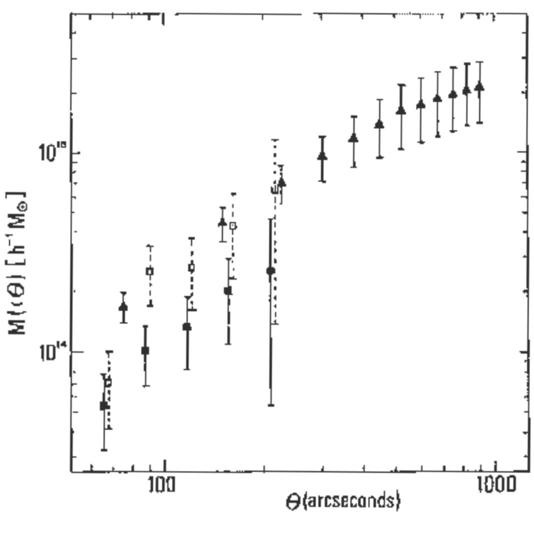

Figure 3.3: The radial mass profiles determined from the X–ray and lensing analysis for Abell 2163. The triangles display the total mass profile determined from the X–ray data. The solid squares are the weak lensing estimates "corrected" for the mean surface density in the control annulus determined from the X–ray data. The conversion from angular to physical units is $60'' = 0.127\,h^{-1}$ Mpc (taken from G. Squires et al., 1997 [29]).

reach out to $2.3\,h^{-1}$ Mpc ($\sim$ 15 *core radii*). The total mass is 2.6 times greater than that of COMA, but the gas mass fraction, $\sim 0.1\ h^{-3/2}$ is typical for rich clusters. A2163 exhibits the Sunyaev–Zel'dovich effect. It is remarkable in the radio, having the most luminous and extended halo yet detected. The galaxy distribution is irregular and extended, with very high velocity dispersions $\sigma = 1680$ km s^{-1}. All data together suggest that there was a recent merger of two large clusters. The optical observations of the distorted images of background galaxies were made with the CFHT telescope. The resulting lensing and X–ray mass profiles are compared in Fig. 3.3. The data sets only overlap out to a radius of $200'' \simeq 500\ h^{-1}$ kpc to which the lensing studies were limited. It is

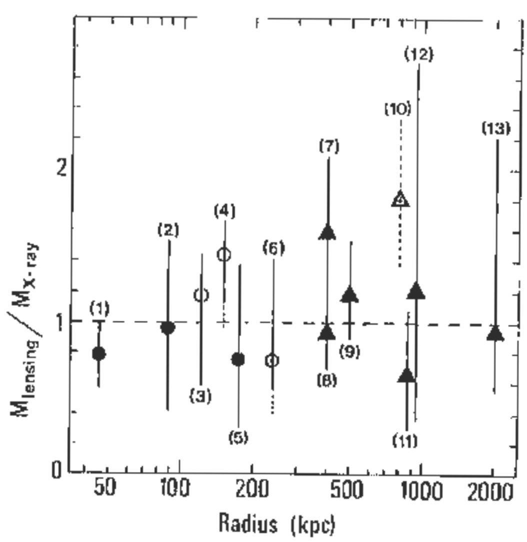

Figure 3.4: The ratio of lensing (strong and weak) and X–ray masses for those clusters for which a reliable and direct comparison of these valnes can be made. The ratios of the strong–lensing and X–ray masses are plotted as circles. Strong lensing resnlts are only presented for the cooling–flow clusters. Filled circles show the results obtained with the detailed lensing models and open circles the results from the simple, spherically–symmetric lensing models (which are only used when results from more detailed modelling are not available). The weak–lensing results are plotted as triangles (taken from S.W. Allen, 1997 [30]).

evident that the lensing mass estimates are systematically lower by a factor of $\simeq 2$ than the X–ray results, but generally the results are consistent with each other, given the substantial uncertainties. There are reasons that the lensing estimate may be biased downward. Correcting for this gives the results displayed by open squares. The agreement between the lensing and X–ray results becomes then quite impressive. The rate and quality of such data will increase dramatically during the coming years. With weak lensing one can also test the dynamical state of clusters. By selecting the relaxed ones one can then determine with some confidence the relative distributions of gas and dark matter. One should also select those cases for applying the Sunyaev–Zel'dovich effect in determining H_0.

In addition, it will become possible to extend the investigations to super clusters scales, with the aim to determine the power spectrum and get information on the cosmological parameters [29]. An interesting comparison of mass measurements for clusters of galaxies using ASCA and ROSAT X–ray data and constraints from strong and weak gravitational lensing has recently been made by S.W. Allen [31]. He showed that for cooling–flow clusters, which are the more dynamically–relaxed systems, the X–ray and strong gravitational lensing mass measurements show excellent agreement, while for the non–cooling–flow clusters, the mass determined from the strong lensing data exceeds the X–ray value by factors $2 \div 4$. On larger spatial scales, comparisons of the X–ray mass results with measurements from weak gravitational lensing show excellent agreement for both cooling–flow and non–cooling–flow clusters.

A summary of these comparisons is shown in Fig. 3.4, which however does not contain the strong lensing data for the non–cooling–flow clusters, since the hydrostatic equilibrium is not expected to hold for these non–relaxed systems. For more information, I refer to the original paper [31] and [32].

Chapter 4

Extensions to a cosmological context

So far, we considered only almost static, weak localized perturbations of Minkowski spacetime. In cosmology the unperturbed spacetime background is given by a Robertson–Walker metric, and this induces various changes in our previous discussions. Fortunately, the final results for the lens map and the time delay look practically unchanged. As it turns out, we only have to insert some obvious redshift factors and interpret all distances as *angular diameter distances*, which is presumably, not really surprising.

4.1 Lens mapping in cosmology

I now describe in more detail the relevant modifications. Let me recall (1.66) (for $c = 1$)

$$\Delta t = \frac{D_d D_s}{2D_{ds}} \left(\frac{\boldsymbol{\xi}}{D_d} - \frac{\boldsymbol{\eta}}{D_s} \right)^2 - \hat{\psi}(\boldsymbol{\xi}) + const.$$

Note that $\left(\frac{\boldsymbol{\xi}}{D_d} - \frac{\boldsymbol{\eta}}{D_s} \right) = (\boldsymbol{\theta} - \boldsymbol{\beta})$. This was the time delay for an almost Newtonian situation. If the distances involved are cosmological, we must multiply the whole expression by the redshift $(1 + z_d)$ of the lens. In addition all distances must be interpreted as angular diameter distances. A systematic derivation is given in the appendix 4.5 to this chapter. Our starting point is thus

$$\Delta t = (1 + z_d) \left\{ \frac{D_d D_s}{2D_{ds}} (\boldsymbol{\theta} - \boldsymbol{\beta})^2 - \hat{\psi}(\boldsymbol{\xi}) \right\} + const. \tag{4.1}$$

The prefactor of the first terms is, clearly, proportional to $1/H_0$ (H_0 is the present Hubble parameter). We shall come back to this.

For cosmological applications, it is convenient to rewrite the potential term slightly. Using the length scale $\xi_0 = D_d$ in (1.65), as well as $\boldsymbol{\theta} = \boldsymbol{\xi}/D_d$, we have

$$\hat{\psi}(\boldsymbol{\xi}) = 4G \int d^2\theta' D_d^2 \Sigma(D_d \boldsymbol{\theta}') \ln |\boldsymbol{\theta} - \boldsymbol{\theta}'| = 2R_s \tilde{\psi}(\boldsymbol{\theta}),$$

where $R_s = 2GM$ is the Schwarzschild radius of the total mass M of the lens, and

$$\tilde{\psi}(\boldsymbol{\theta}) = \int d^2\theta' \tilde{\Sigma}(\boldsymbol{\theta}') \ln |\boldsymbol{\theta} - \boldsymbol{\theta}'|, \tag{4.2}$$

with

$$\tilde{\Sigma}(\boldsymbol{\theta}) := \frac{\Sigma(D_d \boldsymbol{\theta})}{M} D_d^2. \tag{4.3}$$

This quantity gives the fraction of the total mass M per unit solid angle as seen by the observer. We can now write (4.1) in the form

$$\Delta t = \hat{\phi}(\boldsymbol{\theta}, \boldsymbol{\beta}) + const, \tag{4.4}$$

where $\hat{\phi}$ is the *cosmological Fermat potential*

$$\hat{\phi}(\boldsymbol{\theta},\boldsymbol{\beta}) = \frac{1}{2}(1+z_d)\frac{D_d D_s}{D_{ds}}(\boldsymbol{\theta}-\boldsymbol{\beta})^2 - 2R_s(1+z_d)\tilde{\psi}(\boldsymbol{\theta}). \tag{4.5}$$

I should probably stress that in cosmology $D_s \neq D_d + D_{ds}$ (space–time is curved).

It is elementary to work out the angular diameter distance $D(z_1, z_2)$ between two events at red shifts z_1 and z_2 ($z_1 < z_2$). For a Friedmann–Lemaitre model with density parameter Ω_0 and vanishing cosmological constant Λ, one finds

$$D(z_1, z_2) = \frac{1}{H_0\Omega_0^2} 2(1+z_1)\left[R_1(z_2)R_2(z_1) - R_1(z_1)R_2(z_2)\right], \tag{4.6}$$

where

$$R_1(z) = \frac{\Omega_0 z - \Omega_0 + 2}{(1+z)^2}, \qquad R_2(z) = \frac{\sqrt{\Omega_0 z + 1}}{(1+z)^2}. \tag{4.7}$$

The formulas (4.4–7) provide the basis for determinations of the Hubble parameter with gravitational lensing. Some results will be presented later.

From (4.5) we obtain the cosmological lens mapping using Fermat's principle, which implies that $\partial\hat{\phi}/\partial\boldsymbol{\theta} = 0$. This gives

$$\boldsymbol{\beta} = \boldsymbol{\theta} - 2R_s\frac{D_{ds}}{D_d D_s}\frac{\partial\tilde{\psi}}{\partial\boldsymbol{\theta}}. \tag{4.8}$$

For comparison, we write this also in terms of $\boldsymbol{\xi} = D_d\boldsymbol{\theta}, \quad \boldsymbol{\eta} = D_s\boldsymbol{\beta}$ and

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{2R_s}{D_d}\frac{\partial\tilde{\psi}}{\partial\boldsymbol{\theta}} \tag{4.9}$$

as

$$\boldsymbol{\eta} = \frac{D_s}{D_d}\boldsymbol{\xi} - D_{ds}\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}). \tag{4.10}$$

This looks identical to (1.41), but with the present meaning of the symbols it ***holds for arbitrary redshifts.***

Consider two images at the (observed) positions $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$, with separation $\boldsymbol{\theta}_{12} \equiv \boldsymbol{\theta}_1 - \boldsymbol{\theta}_2$ and time delay Δt_{12}. From the lens equation (4.8) we obtain

$$\boldsymbol{\theta}_{12} = 2R_s\frac{D_{ds}}{D_d D_s}\left[\left.\frac{\partial\tilde{\psi}}{\partial\boldsymbol{\theta}}\right|_{\boldsymbol{\theta}_1} - \left.\frac{\partial\tilde{\psi}}{\partial\boldsymbol{\theta}}\right|_{\boldsymbol{\theta}_2}\right]. \tag{4.11}$$

The time delay $\Delta t_{12} = \hat{\phi}(\boldsymbol{\theta}_1,\boldsymbol{\beta}) - \hat{\phi}(\boldsymbol{\theta}_2,\boldsymbol{\beta})$ contains the unobservable angle $\boldsymbol{\beta}$, but this can be eliminated with the lens equation (4.8) and (4.11):

$$\Delta t_{12} = 2R_s(1+z_d)\left\{\frac{1}{2}\left(\left.\frac{\partial\tilde{\psi}}{\partial\boldsymbol{\theta}}\right|_{\boldsymbol{\theta}_1} + \left.\frac{\partial\tilde{\psi}}{\partial\boldsymbol{\theta}}\right|_{\boldsymbol{\theta}_2}\right)\cdot\boldsymbol{\theta}_{12} - \left(\tilde{\psi}(\boldsymbol{\theta}_1) - \tilde{\psi}(\boldsymbol{\theta}_2)\right)\right\}. \tag{4.12}$$

Given a model (i.e. $\tilde{\Sigma}(\boldsymbol{\theta})$), then (4.11) gives a relation

$$\underbrace{\boldsymbol{\theta}_{12}}_{\text{observable}} \leftrightarrow \underbrace{\Omega,\ z_d,\ z_s,\ \frac{2R_s}{H_0^{-1}}}_{\text{dimensionless quantities}},$$

and (4.12) relates Δt_{12} directly to R_s (Ω, H_0, z_s do not appear).

The combination of the two gives R_s and H_0 for given Ω, z_d, z_s. Fortunately, the dependence on Ω is in practice not strong.

Illustration

Consider the simple case of a point source lensed by a point mass (Schwarzschild lens). Then $\tilde{\psi}(\boldsymbol{\theta}) = \ln|\boldsymbol{\theta}|$ and (4.11) gives

$$\theta_{12} = 2R_s \frac{D_{ds}}{D_d D_s}\left(\frac{1}{\theta_1} - \frac{1}{\theta_2}\right),$$

impliying

$$|\theta_1\theta_2| = 2R_s \frac{D_{ds}}{D_d D_s}.$$

On the other hand, equation (4.12) becomes

$$\begin{aligned} \Delta t_{12} &= 2R_s(1+z_d)\left\{\frac{1}{2}\left(\frac{1}{\theta_1}+\frac{1}{\theta_2}\right)\theta_{12} - (\ln|\theta_1| - \ln|\theta_2|)\right\} \\ &= 2R_s(1+z_d)\left\{\frac{\theta_2^2-\theta_1^2}{2|\theta_1\theta_2|} + \ln\left|\frac{\theta_2}{\theta_1}\right|\right\}. \end{aligned}$$

We write this in terms of the ratio ν of the magnifications. Using (2.26) one finds $\nu = \ln(\theta_2/\theta_1)^2$ and thus

$$\Delta t_{12} = R_s(1+z_d)\left\{\nu^{1/2} - \nu^{-1/2} + \ln\nu\right\}.$$

4.2 Hubble constant from time delays

The first term in (4.5) is proportional to H_0^{-1} (see (4.6)). As first noted by Refsdal back in 1964, time delay measurements can yield, in principle, the Hubble parameter. (Note that the lens equation is dimensionless and does thus not provide any constraint.)

Unfortunately, the use of (4.5) *requires a reliable lens model.* This introduces systematic uncertainties. Beside that and H_0, the cosmological Fermat potential involves the density parameter Ω_0 (see (4.6,7)) and Λ (set equal to zero in (4.6) and (4.7)). The dependence on Ω_0 and Λ is, however, not strong, at least in some redshift domains ($z_s \lesssim 2, \quad z_d \lesssim 0.5$).

Table 4.1: Observed and predicted uncertainties in the time delays between images in four gravitationally lensed systems with most secure measured time delays. The best estimate for the H_0 is quoted in the first three cases.

Lens system	$\Delta t/t$ observed	$\Delta t/t$ predicted	best estimate of H_0
Q0957+561	1	10	61
PG1115+080	10	15	53
B0218+357	25	30	70
B1830-211	20	?	?

There is, of course, also the astronomical problem of measuring the time delay. This is not straightforward, as the history of the famous double QSO0957+561 demonstrates. This quasar has been monitored since 1980 both in the optical and radio wavebands, but conflicting claims for Δt have been made. Some of the difficulties are: (*i*) the QSO has not varied strongly, (*ii*) some of the variability is due to microlensing; (*iii*) from the ground the QSO is observable only for 8 months a year at optical telescopes (this does not apply to radio observations, but in practice there are also gaps, as a result of changing configurations of the VLA). Fortunately, the time delay for QSO0957+561 is now well known: $\Delta t = 417 \pm 3$ days [33]. Modelings gave the best estimate, $H_0 \simeq 61$ km s^{-1} Mpc^{-1}. For this example there are constraints for modeling the lens. For example, VLBI images show a core and a radio jet with five blobs, each of which is also doubly imaged. These constraints are really needed, because the lens consists of a galaxy plus a parent cluster and hence requires more parameters for even minimal modeling.[1] It is difficult to assess an error for the value of H_0. Another example is the Einstein ring system B0218+357. A single galaxy is

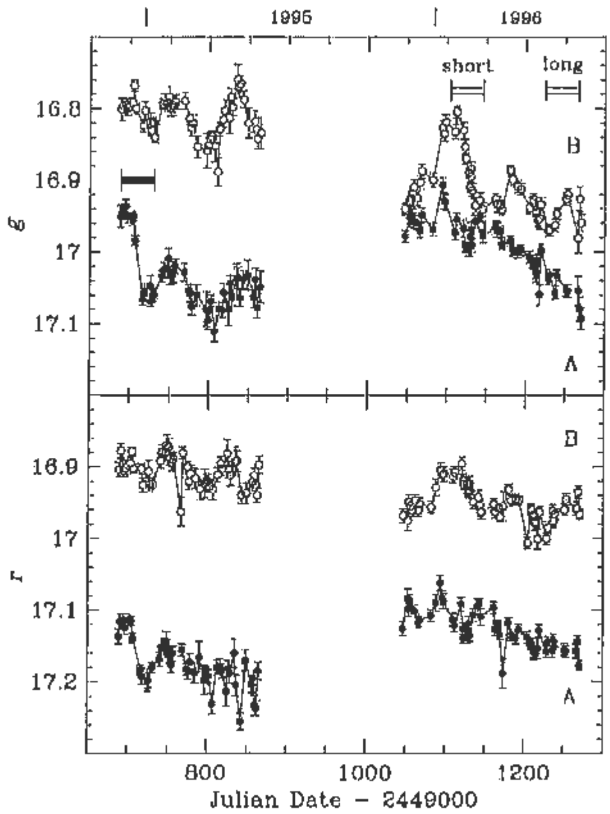

Figure 4.1: Lightcurves of the two images of the gravitationally lensed quasar Q0957+561. Note the sudden decrease of image A at the beginning of the 1995 season (taken from T. Kundić et al., 1997 [31]).

responsible for the small image splitting of 0.3". The time delay was reported to be 12 ± 3 days and the value $H_0 \sim 70$ km s^{-1} Mpc^{-1} was deduced [34]. Further results became known more recently (see Table 4.1; taken from [35].) The ongoing CLASS survey will hopefully uncover new lenses that possess the desirable characteristics for a reliable determination of H_0. Having discussed the problems, I should also mention the advantages of determining H_0 through gravitational lensing over other methods:
(i) The method can be used for *large* redshifts (~ 0.5).
(ii) It is absolutely *independent* of any other method.
(iii) It is based on *fundamental physics*, while other methods rely on models for variable stars (Cepheids), or supernova explosions (type II), or empirical calibrations of standard candles (Tully–Fisher distances, type I supernovae). We repeat, however, that a parametrized lens model is required.

4.3 Bounds on the cosmological constant from lensing statistics

The volnme per unit redshift of the universe at high redshifts increases for a large Λ. This implies that the relative number of lensed sources for a given comoving number density of galaxies increases

[1]Note that the image positions depend on the *derivatives* of the potential, while the potential *itself* determines the time delays. Reconstructing the potential from a small number of derivatives is an ill-posed problem.

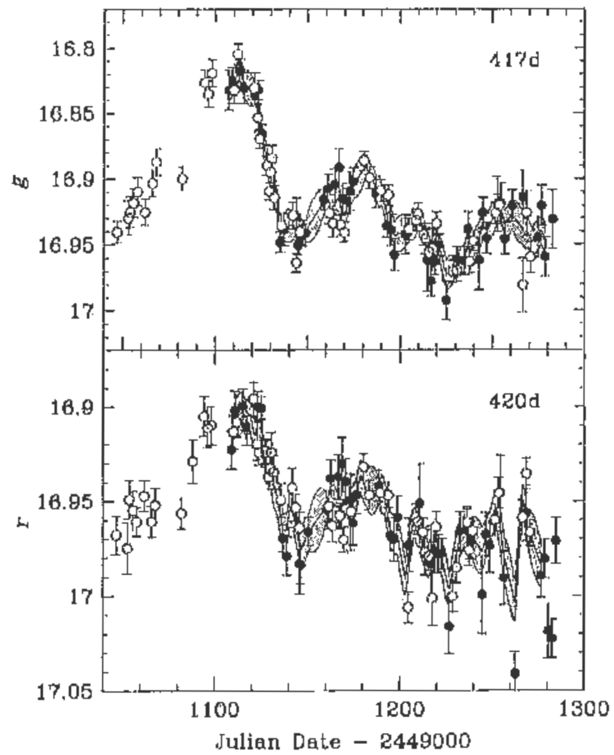

Figure 4.2: The lightcurve of image A of Fig. 4.1 is advanced by the optimal value of the time delay, 417 days (taken from T. Kundić et al., 1997 [31]).

rapidly with Λ. This can be used to constrain Λ by making use of the observed probability of lensing. Various authors have used this method and came up with a limit $\Omega_\Lambda \lesssim 0.6$ for a universe with $\Omega_0 + \Omega_\Lambda = 1$. It remains to be seen whether such bounds, based on lensing statistics, can be improved.

Let me now add a few details. The mentioned volume increase can be seen by looking at the angular distance $D\left(z; \Omega_0, \lambda_0\right)$. For $\Omega_0 + \lambda_0 = 1$ this is given by:

$$H_0 D\left(z\right) = \frac{1}{1+z}\int_0^z \frac{dz'}{\sqrt{\Omega_0\left(1+z'\right)^3 + \left(1-\Omega_0\right)}}. \tag{4.13}$$

In Fig. 4.3, one sees how $D_s(z)$ varies with varying cosmological parameters. For the same four cases, Fig. 4.4 shows the (normalized) probability of a beam encountering a lens for a source at z_s [36]. This is obtained as follows. Let σ be the cross section for "strong" lensing, taken to be πR_E^2, where R_E is the Einstein radius. For a singular isothermal lens this is (see (2.46))

$$\sigma = 16\pi^3\left(\frac{\sigma_v}{c}\right)^4\left(\frac{D_{ds}D_d}{D_s}\right)^2. \tag{4.14}$$

The *optical depth* for a beam of light from a source (z_s) due to lensing is, using $n_d\left(z\right) = n_d\left(0\right)\left(1+z\right)^3 \doteq$ number density of deflectors,

$$\tau\left(z_s\right) = \int_{t_0}^{t_s} n_d c\sigma dt = \int_0^{z_s} n_d\left(0\right)\left(1+z_d\right)^3 \sigma c \frac{dt}{dz_d} dz_d.$$

From the Friedmann Eq. one obtains ($\lambda_0 \equiv \Omega_\Lambda$)

$$cdt = \frac{1}{H_0}\frac{1}{\left(1+z\right)}\frac{dz}{\sqrt{\Omega_0\left(1+z\right)^3 + \left(1-\Omega_0-\lambda_0\right)\left(1+z\right)^2 + \lambda_0}}. \tag{4.15}$$

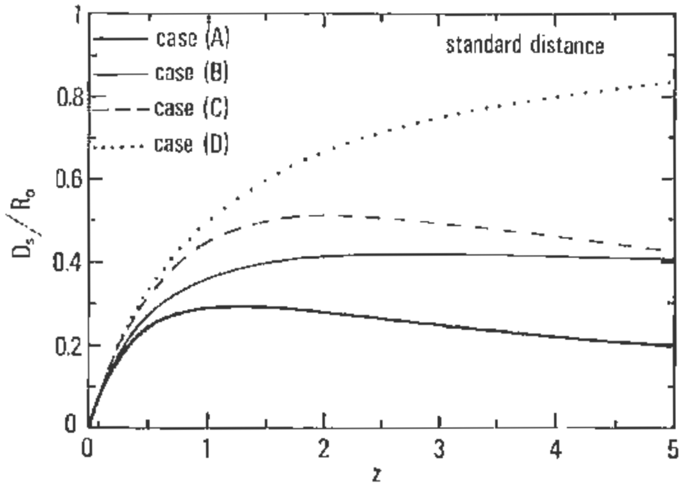

Figure 4.3: Angular diameter distance (D_s) as a function of redshift (z) (in units of $R_0 = c/H_0$). The fonr cases A–D correspond to: $\Omega_0 = 1,\ \Omega_\Lambda = 0$ (A), $\Omega_0 = 0.1,\ \Omega_\Lambda = 0$ (B), $\Omega_0 = 0.1,\ \Omega_\Lambda = 0.9$ (C), $\Omega_0 = 0,\ \Omega_\Lambda = 1$ (D) (taken from M. Fukugita, T. Futamase, M. Kasai and E.L. Turner, 1992 [34]).

Thus

$$\tau(z_s) = n_d(0)\frac{1}{H_0}\int_0^{z_s}\sigma(z_d)\frac{(1+z_d)^2\ dz_d}{\sqrt{\Omega_0(1+z_d)^3+(1-\Omega_0-\lambda_0)(1+z_d)^2+\lambda_0}}. \tag{4.16}$$

Exercise

Equation (4.16) reads

$$\tau(z_s) = n_d(0)\frac{1}{H_0}\int_0^{z_s}\sigma(z)\frac{(1+z)^2}{E(z)}dz,$$

where

$$E^2(z) = \Omega_0(1+z)^3+(1-\Omega_0-\Omega_\Lambda)(1+z)^2+\Omega_\Lambda.$$

Let τ^{fiducial} be the result for $\Omega_0 = 1,\ \Omega_\Lambda = 0$ and show that for a singular isothermal lens

$$\frac{\tau(z_s)}{\tau^{\text{fid}}} = \frac{15}{4}\left[1-\frac{1}{\sqrt{1+z_s}}\right]^{-3}\int_0^{z_s}\frac{(1+z)^2}{E(z)}H_0^2\left(\frac{D_A(0,z)D_A(z,z_s)}{D_A(0,z_s)}\right)^2 dz.$$

The result is plottet in figure 4.4.

Solution

Recall:

$$r(z) = \mathcal{S}(\chi),\ \chi(z) = \frac{1}{a_0H_0}\int_0^z\frac{dz'}{E(z')};\quad D_a(z_1,z_2) = \underbrace{a_0\frac{r_{12}}{1+z_2}}_{a(t_2)r_{12}} = a(t_2)\mathcal{S}(\chi_2-\chi_1).$$

Thus

$$\begin{aligned} D_{ds} &= a(t_s)\mathcal{S}(\chi_s-\chi_d),\\ D_d &= a(t_d)\mathcal{S}(\chi_d),\\ D_s &= a(t_s)\mathcal{S}(\chi_s). \end{aligned}$$

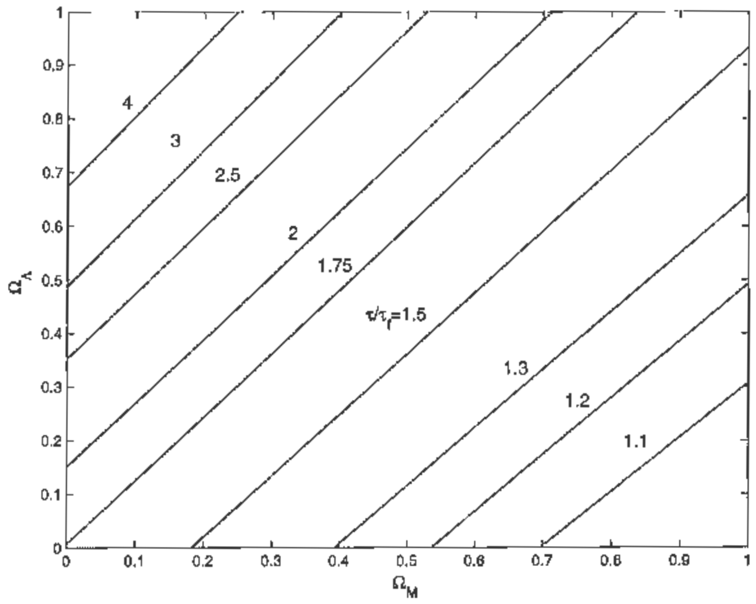

Figure 4.4: Probability for observing a gravitational lens, as contours in the $(\Omega_M, \Omega_\Lambda)$ plane, normalized to unity for the case $\Omega_M = 1,\ \Omega_\Lambda = 0$.

The optical depth τ for a singular isothermal lens (cross section (4.14)) can be written as follows. Start from the equation below (4.14):

$$\tau = \int n_d(0) \underbrace{(1+z_d)^3}_{[a_0/a(t_d)]^3} \sigma \underbrace{dt_d}_{a(t_d)d\chi_d} ,$$

where $\sigma = \pi\hat{\alpha}^2(D_{ds}D_d/D_s)^2$ with $\hat{\alpha} = 4\pi(\sigma_v/c)^2$, to find

$$\tau = \int_0^{\chi_s} n_d(0) a_0^3 \pi\hat{\alpha}^2 \left(\frac{\mathcal{S}(\chi_d)\mathcal{S}(\chi_s - \chi_d)}{\mathcal{S}(\chi_s)}\right)^2 d\chi_d .$$

For the fiducial case $k = 0,\ \Omega_M = 1,\ \Omega_\Lambda = 0$ we have $\mathcal{S}(\chi) = \chi$ and

$$a_0\chi_s = a_0 \frac{D_s}{a(t_s)} = (1+z_s)D_s, \quad D_s = \frac{2}{H_0(1+z_s)}\left[1 - \frac{1}{\sqrt{1+z_s}}\right].$$

Hence,

$$\begin{aligned}
\tau^{\rm fid} &= n_d(0)a_0^3\pi\hat{\alpha}^2 \underbrace{\int_0^{\chi_s}\left[\frac{\chi_d(\chi_s-\chi_d)}{\chi_s}\right]^2 d\chi_d}_{\frac{1}{30}\chi_s^3} \\
&= n_d(0)\pi\hat{\alpha}^2 \frac{1}{30}\left[\frac{2}{H_0}\left(1 - \frac{1}{\sqrt{1+z_s}}\right)\right]^3 \\
&= n_d(0)16\pi^3\left(\frac{\sigma_v}{c}\right)^2 \frac{1}{H_0^3}\frac{4}{15}\left[1 - \frac{1}{\sqrt{1+z_s}}\right]^3 ,
\end{aligned}$$

and with $d\chi = \frac{1}{a_0 H_0}\frac{1}{E(z)}dz$

$$\frac{\tau}{\tau^{\rm fid}} = \frac{15}{4}\left[1 - \frac{1}{\sqrt{1+z_s}}\right]^{-3} \int_0^{z_s} \frac{(1+z_s)^2}{E(z)}\left[\frac{H_0D_A(0,z)H_0D_A(z,z_s)}{H_0D_A(0,z_s)}\right]^2 dz.$$

In Fig. 4.5, the cross section (4.12) has been chosen, in which the $D's$ were taken to be the angular distances ((4.11) for $\Omega_0 + \lambda_0 = 1$). There is actually a problem as to which of the

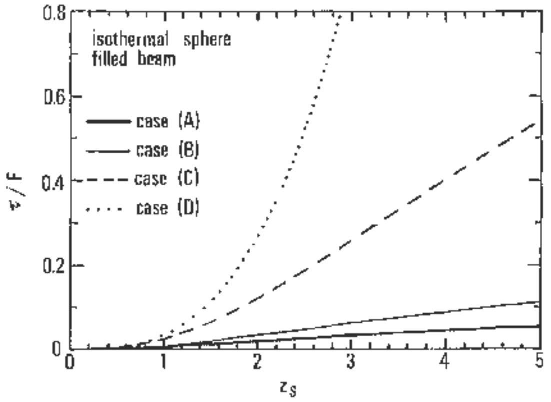

Figure 4.5: Normalized optical depth (see text; taken from M. Fukugita, T. Futamase, M. Kasai and E.L. Turner, 1992 [34]).

redshift–distance relations is the relevant one. This is associated with the fact that the light propagates through an *inhomogeneous* spacetime, rather than the averaged smooth Friedmann–Lemaitre spacetime; the light rays thus feel the local metric. This point has been discussed a lot and causes a significant uncertainty. Various choices have been used in [36]. The main outcome of this investigation was, that the cosmological constant, if it dominates over the mass density, *increases the optical depth greatly* (Fig. 4.4), and that its effect is *much larger* than the uncertainty arising from the choice of the redshift–distance relation.[2] (I guess that Monte–Carlo studies are needed to reduce this inherent uncertainty.) From $\tau(z_s)$ one can predict lensing frequencies if the redshift distribution of quasars is known. In [36] a detailed discussion of the uncertainties, both theoretical and observational, is given. An observational problem is that quasar samples are needed which are homogeneous surveys for lenses. So far, in practice there are biases against lens images with certain $\Delta\theta$.

Another interesting quantity is the *mean image separation* $\overline{\Delta\theta}(z_d)$ at a given z_d, and its average over the lens redshift distribution:

$$< \Delta\theta >= \frac{1}{\tau(z_s)} \int_0^{z_s} \overline{\Delta\theta} \frac{d\tau}{dz_d} dz_d. \tag{4.17}$$

Consider, as an example, again a singular isothermal lens sphere. From the discussion in Section 2.3, we know that the lens produces two images for $\beta < \beta_{crit} = 4\pi\left(\frac{\sigma_v}{c}\right)^2 \frac{D_{ds}}{D_s} = \hat{\alpha}\frac{D_{ds}}{D_s}$. The separation $\Delta x = 2$ translates into

$$\Delta\theta = 2\beta_{crit} = 2\frac{4\pi\sigma_v^2}{c^2}\frac{D_{ds}}{D_s}.$$

Since this is independent of the impact parameter, the right hand side is the value of $\overline{\Delta\theta}$. Thus

$$\frac{\overline{\Delta\theta}}{\hat{\alpha}} = 2\frac{D_{ds}}{D_s}, \qquad \hat{\alpha} = 4\pi\left(\frac{\sigma_v}{c}\right)^2. \tag{4.18}$$

[2]The parameter F in Fig. 4.4 is defined by $F = 16\pi^3 n_d(0)(\sigma_v/c)^4 (c/H_0)^3$ and is a measure for the effectiveness of matter in producing double images.

It turns out [36], that the image separation is relatively insensitive to the choice of the cosmological parameters (and to the choice of the distance formula).

As an exercise, compute $<\Delta\theta>$ for $\Omega_0 + \Omega_\Lambda = 1$. The result is

$$<\Delta\theta>=\hat{\alpha}. \tag{4.19}$$

It has been pointed out by several authors, that the mean separation does become sensitive to Λ, when it is used together with other information (lens redshift, lens magnitude, velocity dispersion of lens galaxy). For a recent paper (with references) see [37]. These authors compare the theoretical prediction of the critical radius θ_{crit} as a function of z_s, z_d, and the apparent magnitude, m_d, of the lens with observations of elliptical (field) galaxies acting as strong gravitational lenses.

For the function θ_{crit} $(z_d,\ z_s,\ m_d)$ they use the singular isothermal model, for which we found

$$\theta_{crit} = 4\pi\left(\frac{\sigma}{c}\right)^2\frac{D_{ds}}{D_s}.$$

In addition they make use of the Faber–Jackson relation

$$\sigma = \sigma_*\left(\frac{L}{L_*}\right)^\beta.$$

Besides this, use the fact that the angular diameter distance D is related to the luminosity distance D_{lum} by $D = D_{lum}/\left(1+z\right)^2$. Introducing now the magnitudes

$$L/L_* = 10^{-\frac{2}{5}(M-M_*)}, \quad m - M = 5\log_{10}\left[D_{lum}\left(Mpc\right)\right] + 25,$$

one gets

$$\theta_{crit} = 4\pi\left(\frac{\sigma_*}{c}\right)^2\frac{D_{ds}}{D_s}\left[D_d\left(1+z_d\right)^2\right]^{4\beta}\cdot 10^{-0.8\beta[m-M_*-25]}. \tag{4.20}$$

To this one has to add the K–correction and an evolutionary correction. Fig. 4.5 below shows the

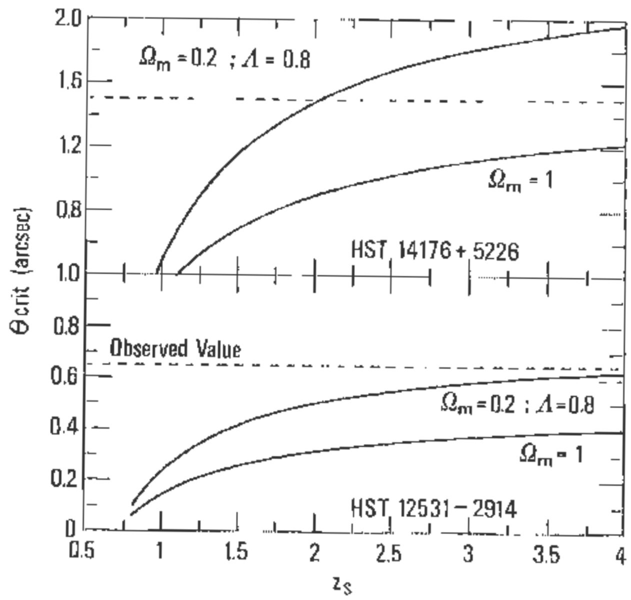

Figure 4.6: $\theta_{crit} - z$ relation for two gravitational lenses HST14156+5226 and HST12531-2914 (see text; taken from M. Im, R.E. Griffiths and K.U. Ratnatunga, 1997 [35]).

$\theta_{crit} - z$ relation for the strong gravitational lens system HST12531-2914 and HST14176+5226, adopting the parameters σ_*, β , etc. described in the paper, and two choices of the cosmological parameters. As one can see, θ_{crit} is quite sensitive to Λ for sufficiently large redshifts z_s. The measured values of θ_{crit} are also shown. Unfortunately, the redshifts of the two sources are not known. There are, however, other good examples where z_s is known. On the basis of a likelyhood analysis of seven strong lenses, satisfying certain selection criteria, the authors come up with the result

$$\Omega_\Lambda = 0.64^{+0.15}_{-0.26}\,, \qquad \text{for} \qquad \Omega_0 + \Omega_\Lambda = 1.$$

Stronger contraints should become possible with future HST observations when new lens systems with measured values of z_d, m_d, z_s are discovered. For a discussion of the intrinsic systematic uncertainties, I refer to the quoted paper.

4.4 Updates

4.4.1 Statistics of strong gravitational lensing of distant quasars by galaxies

I discuss here mainly a recent re-analysis of M. Chiba and Y. Yoshii [38]; see also Yu-N. Cheng and L. M. Krauss [40].

We saw already that the number of multiply imaged QSOs in lens surveys is a sensitive function of Ω_Λ. Observationally there are only a few lenses among hundreds of QSOs.

The re-analysis [38] is based on an improved luminosity function (LF) of E/S0 galaxies and updated knowledge of internal dynamics (velocity dispersions and light profiles). It turns out (as was known before) that spiral and irregular galaxies make negligible contributions to lensing statistics. This may, however, be questioned, since spiral galaxy lenses have been found (e.g. B1600+34 in the CLASS sample).

The lens model used by the authors is an isothermal sphere with finite core,

$$\rho(r) = \frac{\sigma^2}{2\pi G\,(r^2 + r^2_{\mathrm{core}})}, \tag{4.21}$$

that we discussed brieftly in section 2.4 (see also [41]).

For the number density of E/S0 galaxies, the luminosity function ϕ_g in

$$dn_g = \phi_g\left(\frac{L}{L_*}\right)\frac{dL}{L_*} \tag{4.22}$$

is parameterized in the form proposed by Schechter

$$\phi_g(y) = \phi_* y^\alpha e^{-y}. \tag{4.23}$$

Such a Schechter function is also used for the quasar LF ϕ_Q.

We are interested in the probability $p(L_Q, z_s)$ that a QSO with redshift z_s and luminosity L_Q is multiply lensed. This is given by the expression

$$p(L_Q, z_s) = \int_0^{z_s} dz_d (1+z_d)^3 \left|\frac{dt}{dz_d}\right| \int_0^\infty dL_g \phi_g(L_g) \int d\sigma S \frac{\phi_Q(L_Q/\mu, z_s)/\mu}{\phi_Q(L_Q, z_s)}, \tag{4.24}$$

where S is the selection function and the last fraction is the number of QSOs that are amplified to the luminosity L_Q (magnification bias).

The differential cross section is given by $2\pi\ell d\ell$, where $\ell = (D_d/D_s)\eta$ is the impact parameter of the unlensed rays in the lens plane. The integration is restricted to $\ell < \ell_0(L_g, z_s)$, where ℓ_0 is the critical ℓ with 3-fold images for $\ell < \ell_0$.

(**Exercise:** Determine ℓ_0 as a function of $\sigma, r_{\mathrm{core}}$.)

The observational material and fitting procedure entering in (4.24) are described in the original paper [38]. I discuss only some of the main results. Throughout, a flat cosmology ($\Omega_M + \Omega_\Lambda = 1$) is assumed.

The upper plot of fig. 4.7 shows the results of model calculations with different LFs for the adopted surveys (about 900 QSOs at $z_s > 1$ with 5 lensed cases). The three different lines for a particular LF (such as LPEM) correspond to three different values of a faint cutoff magnitude appearing in the parameterization. The lower figure shows the prediction of image-separation, compared with the observational histogram.

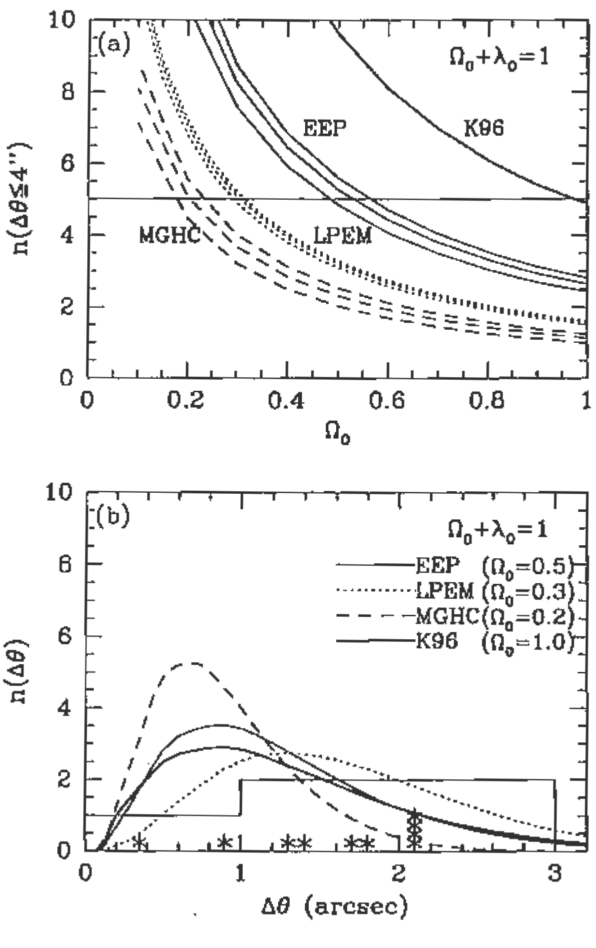

Figure 4.7: (a) Predicted total number of lenses n with $\Delta\theta \leq 4''$ in the adopted optical lens surveys, compared with the observed five lenses (thin solid line). (b) Predicted image-separation distribution $n(\Delta\theta)$, compared with the observed image-separation distribution in the optical sample (histogram) and in the optical lenses (asterisks located at their respective separations $\Delta\theta$).

Fig. 4.8 gives the result of a maximum likelihood analysis for reproducing both the total number of optical lenses n with $\Delta\theta \leq 4''$ and the image separation $n(\Delta\theta)$ of optical and radio lenses.

These results are interesting, systematic uncertainties (galaxy luminosity functions, dark matter velocity dispersions, galaxy core radii) remain. Further observational work is required before reliable values for Ω_M and Ω_Λ can be obtained with this method. At the moment the LF based on the Stromlo-APM survey (LPEM) fits best, and a value $\Omega_M \simeq 0.3$ is favored. This luminosity function has, however, recently been criticized by Kochanek et al [39].

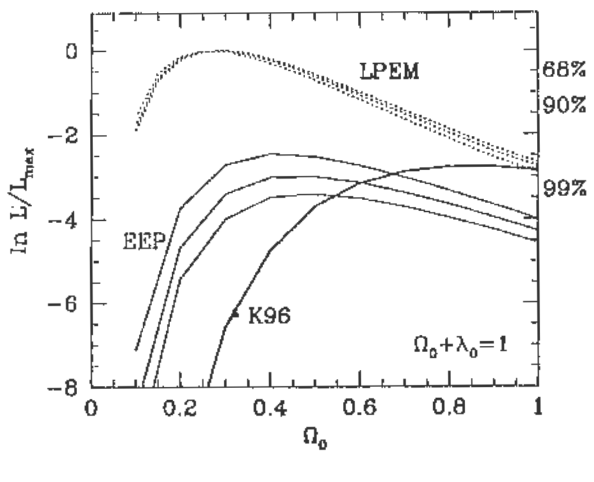

Figure 4.8: Results of the maximul likelihood analysis for reproducing both the total number of optical lenses with $\Delta\theta \leq 4''$ and the image-seperation distribution $n(\Delta\theta)$ of optical and radio lenses.

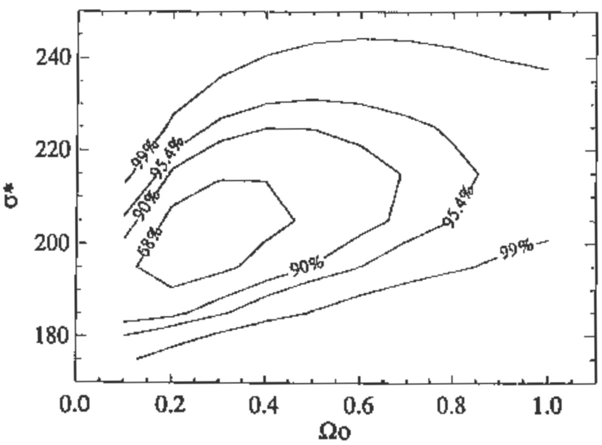

Figure 4.9: Likelihood contour plots for flat cosmologies in the two dimensional parameter space (σ^*, Ω_0), for the standard model using LPEM's LF with $M^{\rm cut} = -17$mag.

4.4.2 Statistics of arcs caused by clusters of galaxies

Clusters with $0.2 \lesssim z_c \lesssim 0.4$ are efficient lenses for background sources at $z_s \sim 1$. For several reasons one can expect that the probability for the formation of pronounced arcs is a sensitive function of Ω_M and Ω_Λ. First, it is well-known that clusters form earlier in low density universes. Secondly, the proper volume per unit redshift is larger for low density universes and depends strongly on Λ for large redshifts (see fig. 6 in [42]).

An extensive numerical study of arc statistics has recently been performed by Bartelmann et al [43], while we have studied this with semi-analytical methods [44, 45].

4.5 Appendix on *Lens mapping in cosmology*

In section 4.1 the basic equation (4.8) was obtained by an educated guess. Below we give an ab initio derivation and also prove the time delay formula (4.4), (4.5).

The null geodesics of the perturbed Friedmann metric (in longitudinal gauge),

$$g = a^2(\eta)\left[-(1+2\phi)d\eta^2 + (1-2\phi)\gamma\right] \tag{4.25}$$

(γ: metric for a space of constant curvature $k = 0, \pm 1$, $\phi \ll 1$), are, after a change of parameter, the same as for the conformally flat metric

$$\tilde{g} = -(1+2\phi)d\eta^2 + (1-2\phi)\gamma \tag{4.26}$$

or

$$\tilde{\tilde{g}} = -d\eta^2 + n^2\gamma, \quad n \simeq 1 - 2\phi. \tag{4.27}$$

Note that we have for $\tilde{g}$ $-(1+2\phi)\dot{\eta}^2 + (1-2\phi)\gamma_{ij}\dot{x}^i\dot{x}^j = 0$ and $(1-2\phi)\dot{\eta}^2 = \text{const}$, thus $\gamma_{ij}\dot{x}^i\dot{x}^j = 1$ for an appropriate normalization of the affine parameter.

The ray orbits $x^i(s)$ thus satisfy the Hamiltonian principle

$$\delta \int n(\mathbf{x}(s))\sqrt{\gamma_{ij}\dot{x}^i(s)\dot{x}^j(s)}ds = 0. \tag{4.28}$$

The corresponding Euler-Lagrange equations are easily found to be

$$\ddot{x}^i + \Gamma^i_{jk}\dot{x}^j\dot{x}^k = (\ln n)^{,i} - (\dot{x}^k\partial_k \ln n)\dot{x}^i, \tag{4.29}$$

where Γ^i_{jk} denote the Crhistoffel symbols for the unperturbed metric γ_{ij}. (The index on the right is raised with γ^{ij}.)

This result also follows from the following **exercise:** Let $\tilde{x}^\mu(\tilde{\lambda})$ be a geodesic for the Riemannian metric $\tilde{g}$ with affine parameter $\tilde{\lambda}$ ($\tilde{g}_{\mu\nu}\frac{dx^\mu}{d\tilde{\lambda}}\frac{dx^\nu}{d\tilde{\lambda}} = 1$):

$$\frac{d^2x^\mu}{d\tilde{\lambda}^2} + \tilde{\Gamma}^\mu_{\alpha\beta}\frac{dx^\alpha}{d\tilde{\lambda}}\frac{dx^\beta}{d\tilde{\lambda}} = 0. \tag{4.30}$$

Consider the conformally related metric g, $\tilde{g} = e^{2\phi}g$, and let λ be a new parameter with $d\tilde{\lambda}/d\lambda = e^{2\phi}$, i.e.,

$$g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = e^{2\phi}. \tag{4.31}$$

1. Show that
$$\frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = e^{2\phi}g^{\mu\nu}\phi_{,\nu} \tag{4.32}$$
($\Gamma^\mu_{\alpha\beta}$: Christoffel symbols for g).
2. Change λ to s, with $g_{\mu\nu}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds} = 1$ and derive the equation
$$\frac{d^2x^\mu}{ds^2} + \Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = g^{\mu\nu}\phi_{,\nu} - \left(\frac{dx^\nu}{ds}\phi_{,\nu}\right)\frac{dx^\mu}{ds}. \tag{4.33}$$

Solution

1. One finds readily
$$\tilde{\Gamma}^\mu_{\alpha\beta} = \Gamma^\mu_{\alpha\beta} + \delta^\mu_\alpha\phi_{,\beta} + \delta^\mu_\beta\phi_{,\alpha} - g_{\alpha\beta}g^{\mu\nu}\phi_{,\nu}.$$
Now,
$$\frac{d^2x^\mu}{d\tilde{\lambda}^2} = e^{-2\phi}\frac{d}{d\lambda}\left(e^{-2\phi}\frac{dx^\mu}{d\lambda}\right) = e^{-4\phi}\left[-2\dot{\phi}\frac{dx^\mu}{d\lambda} + \frac{d^2x^\mu}{d\lambda^2}\right],$$
where $\dot{\phi} \equiv d/d\lambda\phi(x^\mu(\lambda))$. Furthermore,
$$\tilde{\Gamma}^\mu_{\alpha\beta}\frac{dx^\alpha}{d\tilde{\lambda}}\frac{dx^\beta}{d\tilde{\lambda}} = e^{-4\phi}\left\{\Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} + 2\frac{dx^\mu}{d\lambda}\dot{\phi} - \underbrace{g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}}_{e^{2\phi}}g^{\mu\nu}\phi_{,\nu}\right\}.$$
Thus equation (4.32) is correct.

2. We find immediately (4.33) by using

$$\begin{aligned}\frac{dx^\mu}{d\lambda} &= e^\phi \frac{dx^\mu}{ds}\\ \frac{d^2x^\mu}{d\lambda^2} &= e^\phi \frac{d}{ds}\left(e^\phi \frac{dx^\mu}{ds}\right) = e^{2\phi}\left[\frac{d\phi(x(s))}{ds}\frac{dx^\mu}{ds} + \frac{d^2x^\mu}{ds^2}\right].\end{aligned}$$

In abstract notation, the orbit $c(s)$ thus satisfies the equation

$$\nabla_{\dot c}\dot c = \nabla\phi - (\nabla_{\dot c}\phi)\,\dot c \equiv \nabla_\perp \phi. \tag{4.34}$$

$(g(\dot c, \dot c) = 1)$.

For our problem $\tilde g = n^2\gamma$, $g = \gamma$: standard metric of S^3, PS^3, $\mathbb{R}^3$; ∇: covariant derivative for γ;

$$\nabla_{\dot c}\dot c = \nabla_\perp \ln n \qquad \text{(eq. (4.29))}$$

or, with $n = 1 - 2\phi$,

$$\nabla_{\dot c}\dot c = -2\nabla_\perp\phi. \tag{4.35}$$

Consider now a bundle of light rays intersecting at the observer. Each of these rays is characterized by the angles $\boldsymbol\theta$ it encloses with some fiducial ray (see Fig. 4.10). The angles $\boldsymbol\theta$ are all assumed to be small.

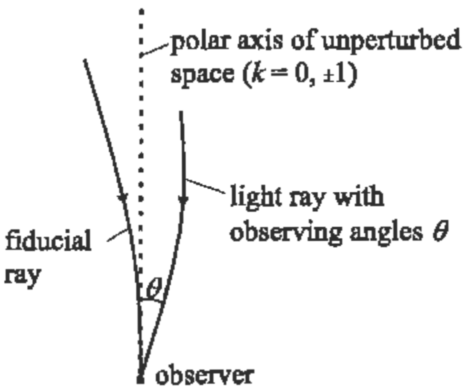

Figure 4.10: Situation under consideration.

The metric γ is one of the standard coodrinates

$$\gamma = d\chi^2 + \mathcal{S}(\chi)d\Omega^2, \qquad \mathcal{S}(\chi) = \begin{cases} \sin\chi & (k=1)\\ \chi & (k=0)\\ \sinh\chi & (k=-1).\end{cases} \tag{4.36}$$

For small angles $\boldsymbol\theta$ we can approximate $d\Omega^2$ of S^2 by $d\theta_1^2 + d\theta_2^2$,

$$\gamma = d\chi^2 + \mathcal{S}^2(\chi)\left(d\theta_1^2 + d\theta_2^2\right). \tag{4.37}$$

In this aproximation, the parameter s satisfies $\dot\chi = 1$, up to quadratic terms. Neglecting such terms also in (4.29) or (4.35), we find

$$\left(\mathcal{S}^2(\chi)\dot\theta_i\right)^\bullet = -2\frac{\partial\phi}{\partial\theta_i}. \tag{4.38}$$

(This is best derived from (4.28).)

It is useful to write these equations also in terms of $x_i \equiv \mathcal{S}(\chi)\theta_i$. We have

$$\begin{aligned}
\dot{x}_i &= \mathcal{S}'(\chi)\dot{\chi}\theta_i + \mathcal{S}\dot{\theta}_i \simeq \mathcal{S}'\chi)\theta_i + \mathcal{S}\dot{\theta}_i, \\
\ddot{x}_i &= 2\mathcal{S}'\dot{\theta}_i + \mathcal{S}''\theta_i + \mathcal{S}\ddot{\theta}_i.
\end{aligned}$$

Since $\mathcal{S}'' + k\mathcal{S} = 0$ we obtain (up to higher orders)

$$\ddot{x}_i + kx_i = \frac{1}{\mathcal{S}}\left(\mathcal{S}^2\dot{\theta}_i\right)^{\bullet} = -2\frac{1}{\mathcal{S}}\frac{\partial\phi}{\partial\theta_i},$$

or

$$\ddot{x}_i + kx_i = -2\frac{\partial\phi}{\partial x_i}. \tag{4.39}$$

We write this in two-dimensional vector notation

$$\frac{d^2\mathbf{x}}{d\chi^2} + k\mathbf{x} = -2\nabla_\perp\phi. \tag{4.40}$$

For $k = 0$ this reduces to the basic equation (1.26). (From this one also obtains immediately the Jacobi equation.)

Now, we proceed as in the appendix 1.8. Since $\mathcal{S}(\chi)$ satisfies

$$\mathcal{S}'' + k\mathcal{S} = 0, \quad \mathcal{S}(0) = 0, \quad \mathcal{S}'(0) = 1, \tag{4.41}$$

the distribution

$$G(\chi, \chi') = \mathcal{S}(\chi)\theta(\chi - \chi') \tag{4.42}$$

is a fundamental solution for the operator $d^2/d\chi^2 + k$ and thus we obtain for $\mathbf{x}(\boldsymbol{\theta}, \chi)$ the integral equation

$$\mathbf{x}(\boldsymbol{\theta}, \chi) = \mathcal{S}(\chi)\boldsymbol{\theta} - 2\int_0^{\chi} \mathcal{S}(\chi - \chi')\nabla_\perp\phi\left(\mathbf{x}(\boldsymbol{\theta}, \chi'), \chi'\right) d\chi'. \tag{4.43}$$

The first term on the right is the unperturbed homogeneous solution of (4.40).

Under the integral we replace $\mathbf{x}(\boldsymbol{\theta}, \chi')$ by the unperturbed solution $\mathcal{S}(\chi')\boldsymbol{\theta}$ (weak lensing)

$$\mathbf{x}(\boldsymbol{\theta}, \chi) = \mathcal{S}(\chi)\boldsymbol{\theta} - 2\int_0^{\chi} \mathcal{S}(\chi - \chi')\nabla_\perp\phi\left(\mathcal{S}(\chi')\boldsymbol{\theta}, \chi'\right) d\chi'. \tag{4.44}$$

Note that this approximate solution satisfies, as it should,

$$\left.\frac{d\mathbf{x}}{d\chi}\right|_{\chi=0} = \boldsymbol{\theta}. \tag{4.45}$$

Let $\boldsymbol{\eta} = \mathbf{x}(\boldsymbol{\theta}, \chi_s)$ for a source at $\chi = \chi_s$. From (4.44) we get

$$\boldsymbol{\eta} = \mathcal{S}(\chi_s)\boldsymbol{\theta} - 2\int_0^{\chi_s} d\chi'\mathcal{S}(\chi - \chi')\nabla_\perp\phi(\text{u.p.})$$

(where 'u.p.' denotes $(\mathcal{S}(\chi)\boldsymbol{\theta}, \chi)$). Setting $\boldsymbol{\eta} = \mathcal{S}(\chi_s)\boldsymbol{\beta}$ ($\boldsymbol{\beta}$: unperturbed position angles of the source), we can write this as

$$\boldsymbol{\beta} = \boldsymbol{\theta} - 2\int_0^{\chi_s} d\chi\frac{\mathcal{S}(\chi_s - \chi)}{\mathcal{S}(\chi_s)}\nabla_\perp\phi(\text{u.p.}). \tag{4.46}$$

The first factor under the integral is often slowly varying over the dimensions of the lens[3] and can thus be replaced by the ratio D_{ds}/D_s of angular diameter distances

$$D_{ds} = a(t_s)\mathcal{S}(\chi_s - \chi_d), \quad \text{etc.}$$

[3]This is not true for weak lensing produced by large scale structures. Then one has to work with (4.46).

In this thin lens approximation we finally obtain

$$\beta = \theta - \frac{D_{ds}}{D_s}\hat{\alpha}, \quad \hat{\alpha} = 2\nabla_\perp \hat{\psi} = \frac{2}{\mathcal{S}(\chi_d)}\frac{\partial \hat{\psi}}{\partial \theta}, \tag{4.47}$$

with

$$\hat{\psi} \equiv \int \phi(\text{u.p.})d\chi. \tag{4.48}$$

This agrees with (4.8), as we now show.

From the last two equations we obtain again

$$\nabla_\perp \cdot \hat{\alpha} = 2\Delta_\perp \hat{\psi} = 2\int \Delta_\perp \phi(\text{u.p.})d\chi = 2\int \underbrace{\Delta\phi(\text{u.p.})}_{4\pi G\rho a^2(t)} d\chi \simeq 8\pi G\Sigma a(t_d),$$

i.e.,

$$\Delta_\perp \hat{\psi} = 4\pi G a(t_d)\Sigma. \tag{4.49}$$

Hence,

$$\begin{aligned} \hat{\psi}(\mathbf{x}) &= 2G\int \ln|\mathbf{x}-\mathbf{x}'| a(t_d)\Sigma(\mathbf{x}') \underbrace{d^2x'}_{\mathcal{S}^2(\chi_d)d^2\theta' = a^{-2}(t_d)D_d^2 d^2\theta'} \\ &= \frac{1}{a(t_d)}2G\int \ln|\theta-\theta'|\Sigma(D_d\theta')d^2\theta' + \text{"const" (indep. of } \theta). \end{aligned}$$

As a result, we finally get

$$\hat{\alpha} = \frac{2}{\mathcal{S}(\chi_d)}\frac{\partial \hat{\psi}}{\partial \theta}(\mathcal{S}(\chi_d)\theta) = \frac{2R_s}{D_d}\frac{\partial \tilde{\psi}}{\partial \theta}, \tag{4.50}$$

where

$$\tilde{\psi}(\theta) = \int \ln|\theta-\theta'|\frac{\Sigma(D_d\theta')}{M}D_d^2 d^2\theta'. \tag{4.51}$$

(4.47), (4.50) and (4.51) agree with the basic formulae (4.26), (4.27), (4.32) and (4.33).

Exercise

Let γ be the standard metric on $M = S^3$, PS^3 or $\mathbb{R}^3$. Let $c: I \to M$ be a geodesic, and X a Jacobi field along c perpendicular to $\dot{c}$. Write the Jacobi equation

$$\ddot{X} + R(X,\dot{c})\dot{c} = 0,$$

where $\dot{X} = \nabla_{\dot{c}}X$, etc., in terms of the components $\xi^i(s)$ of X relative to an orthonormal basis e_1, e_2 orthogonal to $\dot{c}$, which is parallel along c. Result:

$$\ddot{\xi}^i + k\xi^i = 0.$$

Next, we derive the time delay formulas (4.4) and (4.5). For light rays we have $d\eta^2 = n^2\gamma_{ij}dx^i dx^j$, and thus the conformal light travel time is (as in (1.61))

$$\eta = \int_{\text{path}} n\sqrt{\gamma_{ij}\dot{x}^i\dot{x}^j}ds \simeq \ell - 2\int_{\text{u.p.}} \phi d\chi. \tag{4.52}$$

Here ℓ is the path length measured with the unperturbed metric γ_{ij}. We show below that the corresponding conformal time delay, $\Delta\eta_{\text{geom}}$, is given by

$$\Delta\eta_{\text{geom}} = \frac{1}{2}\frac{S(\chi_d)S(\chi_s)}{S(\chi_s-\chi_d)}(\theta-\beta)^2. \tag{4.53}$$

The last integral in (4.52) is $\hat{\psi} = (R_s/a_d)\tilde{\psi}$. Therefore, the total conformal time delay (relative to the unlensed situation) is

$$\Delta\eta = \frac{1}{2}\frac{S(\chi_d)S(\chi_s)}{S(\chi_s - \chi_d)}(\boldsymbol{\theta} - \boldsymbol{\beta})^2 - \frac{2R_s}{a_d}\tilde{\psi}(\boldsymbol{\theta}). \tag{4.54}$$

Since this is small in comparison to the Hubble time H_0^{-1}, we have $\Delta t = a_0\Delta\eta$. Using also the angular diameter distances

$$D_d = a_d S(\chi_d), \quad D_s = a_s S(\chi_s), \quad D_{ds} = a_{ds} S(\chi_s - \chi_d), \tag{4.55}$$

we obtain the important result which we guessed in (4.4) and (4.5):

$$\Delta t = \frac{1}{2}(1 + z_d)\frac{D_d D_s}{D_{ds}}(\boldsymbol{\theta} - \boldsymbol{\beta})^2 - 2R_S(1 + z_d)\tilde{\psi}(\boldsymbol{\theta}). \tag{4.56}$$

Geometrical time delay

We still have to show that for small $\hat{\boldsymbol{\alpha}}$

$$\Delta t_{\text{geom}} = (1 + z_d)\frac{1}{2}\frac{D_d D_s}{D_{ds}}(\boldsymbol{\theta} - \boldsymbol{\beta})^2. \tag{4.57}$$

We give the proof for $k = 0$ in such a way that it can be translated verbatim to $k = \pm 1$ by using the corresponding cosine and sine theorems for spherical and hyperbolic geometries, respectively.

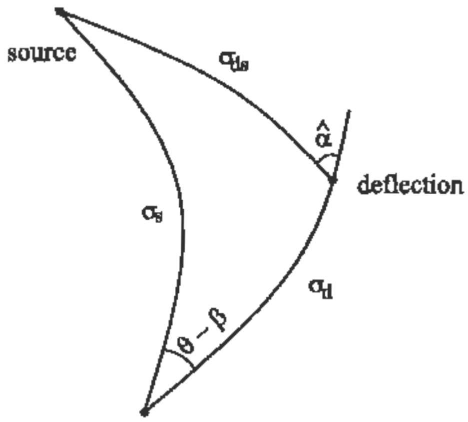

Figure 4.11: The geodesic triangle.

We compute

$$\Delta\eta_{\text{geom}} = \sigma_{ds} + \sigma_d - \sigma_s \tag{4.58}$$

of figure 4.11, and use afterwards $\Delta t_{\text{geom}} = a_0 \Delta\eta_{\text{geom}}$.

The cosine theorem gives

$$\sigma_s^2 = \sigma_d^2 + \sigma_{ds}^2 + 2\sigma_d\sigma_{ds}\cos\hat{\alpha} = (\sigma_d + \sigma_{ds})^2 - 4\sigma_d\sigma_{ds}\sin^2\frac{\hat{\alpha}}{2}.$$

Thus (for small $\hat{\alpha}$)

$$\sigma_s \simeq \sigma_d + \sigma_{ds} - \frac{2\sigma_d\sigma_{ds}}{\sigma_d + \sigma_{ds}}\sin^2\frac{\hat{\alpha}}{2} \simeq \sigma_d + \sigma_{ds} - \frac{1}{2}\frac{\sigma_d\sigma_{ds}}{\sigma_s}\hat{\alpha}^2.$$

Hence,

$$\Delta\eta_{\text{geom}} = \frac{\sigma_d\sigma_{ds}}{2\sigma_s}\hat{\alpha}^2.$$

The sine theorem gives for small $\hat{\alpha}$, $(\boldsymbol{\theta}-\boldsymbol{\beta})$

$$(\boldsymbol{\theta}-\boldsymbol{\beta})\sigma_s = \hat{\boldsymbol{\alpha}}\sigma_{ds} \rightarrow \hat{\boldsymbol{\alpha}}^2 = \frac{\sigma_s^2}{\sigma_{ds}^2}(\boldsymbol{\theta}-\boldsymbol{\beta})^2.$$

Together we obtain

$$\Delta\eta_{\text{geom}} = \frac{\sigma_d\sigma_s}{2\sigma_{ds}}(\boldsymbol{\theta}-\boldsymbol{\beta})^2. \tag{4.59}$$

Now, the angular diameter distances are for $k=0$

$$D_d = a_d\sigma_d, \quad D_s = a_s\sigma_s, \quad D_{ds} = a_s\sigma_{ds}. \tag{4.60}$$

Inserting this gives (with $a_0/a_d = 1+z_d$) indeed (4.57).

Exercise

Translate the argument to the case $k=1$ and show that

$$\Delta\eta_{\text{geom}} = \frac{\sin\sigma_d \sin\sigma_s}{2\sin\sigma_{ds}}(\boldsymbol{\theta}-\boldsymbol{\beta})^2. \tag{4.61}$$

For $k=-1$ one has to replace sin by sinh.

Chapter 5

Complex formulation of lensing theory

This chapter contains parts of my paper [23]. What follows is the abstract.

The elegance and usefulness of a complex formulation of the basic lensing equations is demonstrated with a number of applications. Using standard tools of complex function theory, we present, for instance, a new proof of the fact that the number of images produced by a regular lens is always odd, provided that the source is not located on a caustic. Several differential and integral relations between the mean curvature and the (reduced) shear are also derived. These emerge almost automatically from complex differentiations of the differential of the lens map, together with Stokes' theorem for complex valued 1–forms.

5.1 Complex formulation

In this section we translate the basic lensing equations into a complex formulation. It will turn out that this is not only elegant, but also quite useful, because one can then apply various tools and techniques of complex analysis. This has also been noted before by other authors.

5.1.1 Mathematical preliminaries

We use standard notation when identifying $\mathbb{R}^2$ with $\mathbb{C}$, by writing $z = x + iy$ for $(x, y) \in \mathbb{R}^2$ and $dz = dx + idy$, $d\bar{z} = dx - idy$ for the corresponding basis of 1–forms. In terms of the Wirtinger derivatives,

$$\partial_z \equiv \frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right), \quad \partial_{\bar{z}} \equiv \frac{\partial}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right), \tag{5.1}$$

the differential of any smooth complex function f on $\mathbb{C}$ has the representation

$$df = \frac{\partial f}{\partial z}dz + \frac{\partial f}{\partial \bar{z}}d\bar{z}. \tag{5.2}$$

We shall also write f_z and $f_{\bar{z}}$ for $\partial_z f$ and $\partial_{\bar{z}} f$, respectively. A function f is holomorphic if and only if $\partial_{\bar{z}} f = 0$. In terms of the Wirtinger derivatives, the Laplacian is given by

$$\Delta = 4\partial_z\partial_{\bar{z}}. \tag{5.3}$$

We shall make repeated use of Stokes' theorem for complex–valued differential forms on $\mathbb{C}$ (or an open subset of $\mathbb{C}$): If Ω is a compact subset of $\mathbb{C}$ with a smooth boundary $\partial\Omega$, then for every complex differential 1–form ω

$$\int_\Omega d\omega = \int_{\partial\Omega} \omega. \tag{5.4}$$

An immediate corollary of Eq. (5.4) is the Cauchy–Green formula: For a smooth function f we consider

$$\omega = f\frac{dz}{z - \zeta}, \tag{5.5}$$

and apply Stokes' theorem (5.4) for Ω minus an ε–disk with center ζ. In the limit $\varepsilon \to 0$ we obtain

$$f(\zeta) = \frac{1}{2\pi i}\int_{\partial\Omega}\frac{f(z)}{z-\zeta}dz + \frac{1}{2\pi i}\int_{\Omega}\frac{f_{\bar z}(z)}{z-\zeta}dz\wedge d\bar z. \tag{5.6}$$

For holomorphic functions the second integral is absent. (Note that $dz \wedge d\bar z = -2idx \wedge dy$.)

The *dilatation* or *Beltrami coefficient* $\nu = \nu_f$ of a smooth function f is defined by

$$f_{\bar z} = \nu_f f_z, \tag{5.7}$$

and this equation is also called *Beltrami equation*. Since the Jacobian J_f of f is given by

$$J_f = |f_z|^2 - |f_{\bar z}|^2, \tag{5.8}$$

we conclude that $|\nu_f| < 1$ if f preserves orientation and $\nu_f = 0$ if and only if f is conformal. For the interpretation of ν_f we consider the infinitesimal ellipse field by assigning to each $z \in \mathbb{C}$ the ellipse that is mapped to a circle by f. As indicated in Fig. 5.1, the argument of the major axis of this infinitesimal ellipse is $[\pi + \arg(\nu_f)]/2$, and the eccentricity ε is

$$\varepsilon = \frac{|f_z| - |f_{\bar z}|}{|f_z| + |f_{\bar z}|} = \frac{1-|\nu_f|}{1+|\nu_f|}. \tag{5.9}$$

Solving the Beltrami equation (5.7) is then equivalent to finding a function f whose associated ellipse field coincides with a prescribed ν. We shall see that this is just the inversion problem in gravitational lensing. Weak gravitational lensing corresponds to quasiconformal maps. A smooth map f is *k–conformal* if its Beltrami parameter ν_f satisfies $|\nu_f| \le k < 1$. Geometrically, this means that there is a fixed bound on the stretching of f in any given direction compared to any other direction.

We now quote an existence and uniqueness theorem for the Beltrami equation. For a fixed k with $0 < k < 1$ let $L^\infty(k, R)$ denote the measurable functions on $\mathbb{C}$ bounded by k and supported in $\{z \in \mathbb{C} \,|\, |z| < R\}$.

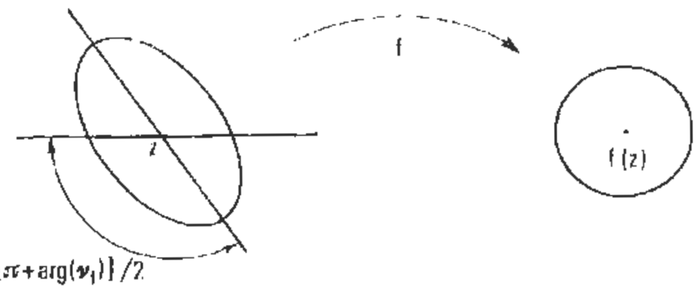

Figure 5.1: Geometrical interpretation of the Beltrami parameter.

Theorem: *For $\nu \in L^\infty(z, R)$, there is a complex function f on $\mathbb{C}$, normalized so that $f(z) = z + O(1/z)$ at ∞, with distributional derivatives satisfying the Beltrami equation $f_{\bar z} = \nu f_z$, and such that $f_{\bar z}$ and $f_z - 1$ belong to L^p for a $p > 2$ sufficiently close to 2. Any such f is unique. The solution f is a homeomorphism of $\mathbb{C}$, which is holomorphic on any open set on which $\nu = 0$. If $\nu \in C^1$ and $\nu_z \in C^1$, then $f \in C^1$.*

A proof of this theorem can, for instance, be found in [46].

The reconstruction problem (for noncritical lensing) will lead to the inhomogeneous Cauchy–Riemann equation

$$\partial_{\bar z} f = h. \tag{5.10}$$

In case the smooth function h has compact support, the Cauchy–Green formula (5.6) provides one solution:

$$f(\zeta) = \frac{1}{2\pi i} \int_{\mathbb{C}} \frac{h(z)}{z-\zeta} dz \wedge d\bar{z}. \tag{5.11}$$

Obviously, f is only determined up to an additive holomorphic function. If the solution is assumed to be bounded, f is *unique up to an additive constant.*

From the solution (5.11) we see that $(\pi z)^{-1}$ is a fundamental solution of the differential operator $\partial_{\bar{z}}$,

$$\frac{1}{\pi} \partial_{\bar{z}} \left(\frac{1}{z} \right) = \delta, \tag{5.12}$$

because (5.11) can be written as

$$f = \frac{1}{\pi z} * h. \tag{5.13}$$

A special case of the so–called Dolbaut Lemma in several complex variables implies that one may drop the assumption that h has compact support:

Theorem: *For any smooth function h on $\mathbb{C}$ there exists a smooth function f such that (5.10) holds.*

For a complete proof, see Chapter 2 of [47].

As an easy consequence we have the

Corollary: *For any smooth function h there exists a smooth solution of the Poisson equation $\Delta f = h$.*

In the following we often use the abbreviations $\partial \equiv \partial_z, \quad \bar{\partial} \equiv \partial_{\bar{z}}$.

5.1.2 The complex lens mapping and its differential

The lens mapping $\varphi : \mathbb{R}^2 \longmapsto \mathbb{R}^2$

$$\mathbf{y} = \varphi(\mathbf{x}) = \mathbf{x} - \nabla \psi(\mathbf{x}), \tag{5.14}$$

is now written as $f : \mathbb{C} \longmapsto \mathbb{C}, \ w = f(z)$ with $z = x_1 + ix_2, \ w = y_1 + iy_2$. We have

$$f(z) = z - 2\bar{\partial}\psi \tag{5.15}$$

or

$$f = \bar{\partial}(z\bar{z} - 2\psi). \tag{5.16}$$

Eq. (1.49) becomes

$$2\partial\bar{\partial}\psi = \kappa. \tag{5.17}$$

The differential of f will be very important. From (5.15) and (5.17) we obtain

$$df = (1-\kappa)\, dz - 2\bar{\partial}^2 \psi d\bar{z}.$$

But

$$\bar{\partial}^2 \psi = \frac{1}{4} \left(\partial_1^2 - \partial_2^2 \right) \psi + \frac{i}{2} \partial_1 \partial_2 \psi = \frac{1}{2} (\gamma_1 + i\gamma_2),$$

according to the original definition (1.52) of the shear vector. Introducing the complex shear

$$\gamma = \gamma_1 + i\gamma_2 \tag{5.18}$$

we obtain

$$df = (1-\kappa)\, dz - \gamma d\bar{z}. \tag{5.19}$$

Hence, the Beltrami parameter ν_f of the lens map is given by

$$\nu_f = -\frac{\gamma}{1-\kappa}. \tag{5.20}$$

This agrees with the reduced shear introduced by Schneider and Seitz [48]. The examples (2.23) and (2.45) become

$$\textit{Schwarzschild lens}: \qquad f(z) = z - \frac{1}{\bar{z}},\ \nu_f = \frac{1}{\bar{z}^2} \tag{5.21}$$

$$\textit{singular isothermal lens}: \qquad f(z) = z - \frac{z}{|\bar{z}|},\ \nu_f = \frac{1}{2}\frac{z}{\bar{z}\left(|z| - \frac{1}{2}\right)}. \tag{5.22}$$

For reference, we note that, according to (1.54), (5.8) and (5.19) the amplification μ is given by

$$\mu^{-1} = |J_f| = \left||\partial f|^2 - |\bar{\partial} f|^2\right| = \left|(1-\kappa)^2 - |\gamma|^2\right|. \tag{5.23}$$

5.2 Applications

The usefulness of the complex formulation will be illustrated in this section with several applications. No new results are obtained, but some of the derivations become simpler and more natural.

5.2.1 Number of images for a regular lens

The important fact that the number of images for a regular lens is always odd, provided the source does not lie on a caustic, is traditionally proven with the help of some elements of Morse theory [5]. We now give a proof which uses only standard tools of complex function theory that are used, for example, in the derivation of the theorem of residues. In particular, we make use of the following analytic formula for the index of a closed (rectifiable) curve γ relative to a point $a \notin \gamma$:

$$ind_\gamma(a) = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z-a}. \tag{5.24}$$

This index is equal to the winding number of γ around a and hence an integer. Furthermore, it is a homotopic invariant, changes sign under orientation reversion, and is additive under composition of closed curves (see, e.g., Chapter IV of [49]).

Consider now a point w_0 in the source plane with images $f^{-1}(w_0) = \{z_1, ...z_N\}$ in the lens plane. The complex 1–form

$$\omega = \frac{1}{2\pi i}\frac{df}{f - w_0} \tag{5.25}$$

is regular on $\mathbb{C}\backslash\bigcup_j D_\varepsilon(z_j)$, where $D_\varepsilon(a)$ denotes the closed disk with center a and radius ε. It is also closed, and therefore Stokes' theorem (5.4) gives

$$\frac{1}{2\pi i}\int_{\partial D_R(0)} \frac{df}{f - w_0} = \sum_{j=1}^{N} \frac{1}{2\pi i}\int_{\partial D_\varepsilon(z_j)} \frac{df}{f - w_0}. \tag{5.26}$$

Now, for a closed curve γ we have by the transformation formula of integrals and (5.24)

$$\frac{1}{2\pi i}\int_\gamma \frac{df}{f - w_0} = \frac{1}{2\pi i}\int_{f\circ\gamma} \frac{dw}{w - w_0} = ind_{f\circ\gamma}(w_0). \tag{5.27}$$

Asymptotically the lens map approaches the identity, and hence the left hand side of (5.26) is equal to 1 for sufficiently large R. Therefore, we have

$$1 = \sum_{j=1}^{N} ind_{f\circ\partial D_\varepsilon(z_j)}(w_0) = n_1 - n_{-1} + 2(n_2 - n_{-2}) + ..., \tag{5.28}$$

where n_λ denotes the number of z_j in $\{z_1, ..., z_N\}$ for which the index in (5.28) is equal to λ.

For the special case, when w_0 is not on a caustic, the Jacobians $J_f(z_j)$ do not vanish and all indices are thus equal to ± 1 (+1 if f is orientation preserving and -1 if it is orientation reversing at z_j). Hence

$$N = n_1 + n_{-1}, \qquad 1 = n_1 - n_{-1}, \tag{5.29}$$

implying that

$$N = 1 + 2n_{-1} \tag{5.30}$$

is odd.

5.2.2 Relations between mean convergence and reduced shear

The Beltrami parameter (reduced shear) ν_f of a lens map is in principle observable. What we are really interested in is, however, the mean curvature κ which is related to the surface mass density by (1.44).

In view of (5.18) it is natural to look first for relations between the complex shear γ and κ.

Eq. (5.19) for the differential of the complex lens map and (5.15) give

$$\gamma = -\bar{\partial} f = 2\bar{\partial}^2 \psi. \tag{5.31}$$

In order to get a useful relation we differentiate (5.31) and use (5.17)

$$\partial\gamma = 2\bar{\partial}\left(\partial\bar{\partial}\psi\right) = \bar{\partial}\kappa. \tag{5.32}$$

This can be regarded as an inhomogeneous Cauchy–Riemann equation for κ. With the results in Subsection 5.2.1 we conclude

$$\kappa = \frac{1}{\pi}\left(\frac{1}{z}\right) * \partial\gamma + \kappa_0 = \frac{1}{\pi}\partial\left(\frac{1}{z}\right) * \gamma + \kappa_0$$

or

$$\kappa = -\frac{1}{\pi}\frac{1}{z^2} * \gamma + \kappa_0. \tag{5.33}$$

The additive constant κ_0 reflects the fact that a homogeneous mass sheet does not produce any shear ("mass sheet degeneracy"). The real form of (5.33) appeared the first time in [20]. In making use of (5.20), we obtain an integral equation for κ when ν is known

$$\kappa = -\frac{1}{\pi}\frac{1}{z^2} * [\nu(1-\kappa)] + \kappa_0. \tag{5.34}$$

This has been used, for instance, in [22] for nonlinear cluster inversions.

We add that (5.34) has an inverse, that also appeared in the influential paper [20] of Kaiser and Squires. From (5.31), (1.48) we obtain

$$\gamma = 4\bar{\partial}^2 \mathcal{G} * \kappa. \tag{5.35}$$

Since the fundamental solution $\mathcal{G}$ of the two–dimensional Laplace operator is

$$\mathcal{G} = \frac{1}{2\pi}\ln|z| = \frac{1}{4\pi}\ln(z\bar{z}), \tag{5.36}$$

we find

$$\gamma = -\frac{1}{\pi}\frac{1}{z^2} * \kappa. \tag{5.37}$$

Note that (5.32) has the real form (κ is real)

$$\nabla\kappa = \begin{pmatrix} \partial_1\gamma_1 + \partial_2\gamma_2 \\ \partial_1\gamma_1 + \partial_2\gamma_1 \end{pmatrix}. \tag{5.38}$$

Let us differentiate (5.32) once more

$$\partial\partial\kappa = \partial^2\gamma, \tag{5.39}$$

giving

$$\Delta\kappa = 4\partial^2\left[\nu\left(1-\kappa\right)\right], \tag{5.40}$$

from where we could again arrive at (5.34). The mass–sheet degeneracy is reflected in the following invariance property: Eq. (5.40), for given ν, remains invariant under the substitution

$$\kappa \to \lambda\kappa + (1-\lambda), \tag{5.41}$$

where λ is a real constant [26].

We can use (5.32) in a different manner. First, we write this equation as

$$\partial\kappa = \partial\left[\nu\left(1-\kappa\right)\right] = (1-\kappa)\,\partial\nu - \nu\partial\kappa.$$

This becomes simpler in terms of $K := \ln(1-\kappa)$

$$\partial K - \nu\partial K = \partial\nu. \tag{5.42}$$

To this we add its complex conjugate. Noting that K is real, we again obtain an inhomogeneous Cauchy–Riemann equation, this time for K:

$$\bar{\partial}K = h(\nu), \tag{5.43}$$

whereby the inhomogeneity

$$h(\nu) = \left(1-|\nu|^2\right)^{-1}\left[\partial\nu + \overline{\nu\partial\nu}\right] \tag{5.44}$$

is, in principle, observable.

The real form of this equation was obtained by Kaiser [24] and has often been used in the analysis of cluster data. The complex version appears also in [50].

The utility of the complex formulation should now be clear. The relations, derived in this subsection, emerge almost automatically by just applying ∂ and ∂ to the coefficients of the differential of the lens map.

5.2.3 Other useful reconstruction equations

Real lensing data are always confined to a finite field of the sky. Therefore, the solution of (5.43) in the form (5.11), for example, involving an integration over all of $\mathbb{C}$, is not very practical. On the other hand, one can obtain integral formulas in which only integrations over bounded domains occur. In order to arrive at the latter, we write the inhomogeneous Cauchy–Riemann equation in terms of differential forms:

$$d''g = \omega. \tag{5.45}$$

Here ω is a 1–form and we use the standard decomposition $d = d' + d''$ of the exterior derivative, satisfying

$$d' \circ d' = 0, \quad d'' \circ d'' = 0, \quad d' \circ d'' + d'' \circ d' = 0 \tag{5.46}$$

(see, e.g., [47]). We also make use of the $*$–operator, which is related to complex conjugations as follows: If a 1–form α is decomposed as $\alpha = \alpha_1 + \alpha_2$, where α_1 is of type (1, 0) and α_2 of type (0, 1), then

$$*\alpha = i\left(\bar{\alpha}_1 - \bar{\alpha}_2\right). \tag{5.47}$$

The following identities are useful

$$**\alpha = -\alpha, \quad \overline{*\alpha} = *\bar{\alpha},$$

$$d*(\alpha_1+\alpha_2) = id'\bar{\alpha}_1 - id''\bar{\alpha}_2,$$

$$*d'g = id''\bar{g}, \quad *d''g = -id'\bar{g},$$

$$d * dg = 2id'd''\bar{g} = \Delta g dx \wedge dy, \tag{5.48}$$

where g is a function.

Now let $\Omega \subset \mathbb{C}$ be a bounded domain with smooth boundary $\partial\Omega$ and $A = |\Omega|$. We show that g minus its average $\bar{g}$ over Ω,

$$\bar{g} = \frac{1}{A}\int_\Omega g dx \wedge dy, \tag{5.49}$$

can be represented in the following form

$$g - \bar{g} = \int_\Omega *\alpha \wedge \omega. \tag{5.50}$$

The 1–form α in the integral is given by

$$\alpha = -2d''H \tag{5.51}$$

in terms of the real Green's function H, defined by

$$\Delta H - \frac{1}{A} = -\delta, \tag{5.52}$$

together with the Neumann boundary condition on $\partial\Omega$.

This is a consequence of Stokes' theorem. The integrand in (5.50) is

$$*\alpha \wedge \omega = *\alpha \wedge d''g = -d''(g * \alpha) - 2gd''(*d''H).$$

By making use of (5.48) we obtain for the last term

$$2gd''(*d''H) = -2igd''d'H = g\Delta H dx \wedge dy,$$

while the first term is given by

$$d''(g * d''H) = d(g * d''H).$$

Hence,

$$\int *\alpha \wedge \omega = \int_{\partial\Omega} g * d''H + g - \bar{g}.$$

This is just (5.50) since the last integral vanishes, due to the Neumann boundary condition for H. Formulas equivalent to (5.50) have been used often by S. Seitz and P. Schneider [22].

The starting point for the derivation of another useful relation is (5.19) in the form

$$d(f - z) = -\kappa dz - \gamma d\bar{z},$$

If we wedge this with $d\bar{z}$ and subtract the complex conjugate of the resulting equation, we find

$$\kappa dz \wedge d\bar{z} = \frac{1}{2}d\left[\kappa(zd\bar{z} - \bar{z}dz) - \gamma\bar{z}d\bar{z} + \bar{\gamma}zdz\right]. \tag{5.53}$$

Taking the average according to (5.49) we arrive at

$$\bar{\kappa} = \langle\kappa\rangle - \frac{\oint(\gamma\bar{z}d\bar{z} - \bar{\gamma}zdz)}{\oint(zd\bar{z} - \bar{z}dz)}, \tag{5.54}$$

where $\langle\cdot\rangle$ denotes the average along the boundary $\partial\Omega$:

$$\langle\kappa\rangle = \frac{\oint\kappa(zd\bar{z} - \bar{z}dz)}{\oint(zd\bar{z} - \bar{z}dz)}. \tag{5.55}$$

For the special case of a disk D_r, we have along the boundary $z = re^{i\varphi}$ $zd\bar{z} - \bar{z}dz = -2ir^2d\varphi$, hence

$$\bar{\kappa} = \langle\kappa\rangle - \langle\gamma_t\rangle, \tag{5.56}$$

where γ_t denotes the tangential component of the shear

$$\gamma_t = \gamma_1 \cos 2\varphi + \gamma_2 \sin 2\varphi. \tag{5.57}$$

This relation is not new (see Ref. [21]). Noting that

$$\bar{\kappa} = \frac{1}{\pi r^2} \int_0^r \kappa \left(r', \varphi\right) r' dr' d\varphi, \tag{5.58}$$

and thus

$$\frac{d\bar{\kappa}}{d \ln r} = 2 \langle \kappa \rangle - 2\bar{\kappa}, \tag{5.59}$$

we can use (5.56) to obtain the interesting connection

$$\frac{d\bar{\kappa}}{d \ln r} = 2 \langle \gamma_t \rangle . \tag{5.60}$$

This has recently been used in an analysis of weak lensing data [20]. A useful integral form of it is, in obvious notation,

$$\bar{\kappa}(r_1) - \bar{\kappa}(r_1 < r < r_2) = -2 \left(1 - \frac{r_1^2}{r_2^2}\right)^{-1} \int_{r_1}^{r_2} \langle \gamma_t \rangle \frac{dr}{r}. \tag{5.61}$$

The left hand side of this equation is what Kaiser and Squires call the ζ–statistics, $\zeta\left(r_1,\ r_2\right)$. One can use general weight functions for the average process [21] and try to optimize the choice for the detection of mass overdensities [22]. Note also, that the integral on the right in (5.61) can be written as

$$\int_{r_1}^{r_2} \langle \gamma_t \rangle \frac{dr}{r} = \frac{1}{2\pi} \int_{[r_1, r_2]} \Re \left(\frac{1}{\bar{z}^2} \bar{\gamma}.\right) dx \wedge dy \tag{5.62}$$

We conclude by pointing out another appearance of a Beltrami parameter in lensing theory. An often used method for describing the shape of a galaxy image uses the second brightness moments

$$Q_{ij} = \frac{1}{\text{Norm}} \int I(\mathrm{x}) \left(x_i - \bar{x}_i\right) \left(x_j - \bar{x}_j\right) d^2x, \tag{5.63}$$

where $I(\mathrm{x})$ is the surface brightness distribution and $\bar{\mathrm{x}}$ is the center of light of the galaxy image. Now regard $Q = (Q_{ij})$ as a linear map of $\mathbb{R}^2$. If this is interpreted as a map $z \longmapsto w(z)$ of $\mathbb{C}$ it reads

$$w = \frac{1}{2} \left(Q_{11} + Q_{22}\right) z + \frac{1}{2} \left(Q_{11} - Q_{22} + 2iQ_{12}\right) \bar{z} = \frac{1}{2} trQ \left[z + \chi \bar{z}\right], \tag{5.64}$$

where

$$\chi = \frac{Q_{11} - Q_{22} + 2iQ_{12}}{\mathrm{tr}Q}. \tag{5.65}$$

χ is called the complex ellipticity and is clearly just the Beltrami parameter of the map (5.64). The intrinsic brightness moments $Q_{ij}^{(s)}$ of the galaxy are defined correspondingly and it is easy to see that $Q^{(s)} = D\varphi \cdot Q \cdot D\varphi, D\varphi$ being the differential (1.52) of the lens map. The interpretation of χ given above, allows us to easily find the corresponding relation between χ and $\chi^{(s)}$. One just has to compose the map (5.64) on the right and on the left with the linearized lens map

$$w = (1 - \kappa) z - \gamma \bar{z}. \tag{5.66}$$

This readily gives

$$\chi^{(s)} = \frac{-2\nu + \chi + \nu^2 \bar{\chi}}{1 + |\nu|^2 - 2\Re(\nu \bar{\chi})}, \tag{5.67}$$

with the inverse

$$\chi = \frac{2\nu + \chi^{(s)} + \nu^2 \chi^{(s)}}{1 + |\nu|^2 + 2\Re\left(\nu \bar{\chi}^{(s)}\right)}. \tag{5.68}$$

A real derivation of these formulas is quite akward. They are used in applications by averaging over a set of galaxy images, together with statistical assumptions about the intrinsic ellipticity distribution (for instance $\langle \chi^{(s)} \rangle = 0$), to determine the reduced shear ν of the lens map. Here, we just wanted to point out that χ has the interpretation of a Beltrami parameter, and that the relations (5.67) and (5.68) are very easily obtained in the complex formalism.

We hope that the reader will find other examples of such simplifications. After this paper was made public, I learnt more about the related work of T. Schramm. As a supplement to what was discussed above, I refer especially to his study of the Beltrami equation with the help of the corresponding characteristic equations [51].

Acknowledgements
The author is very grateful to S. Capozziello, L. Grenacher, G. Iovane, A.A. Marino and M. Kunz for their help in preparing these lecture notes.

Part II

Lectures on Anisotropies in the Cosmic Microwave Background

Chapter 6

Introduction

In these lectures I would like to show you the importance and the power of measurements of anisotropies in the cosmic microwave background (CMB).

CMB anisotropies are so useful mainly because they are small: For a given model, they can be calculated within linear perturbation theory, to very good approximation. They are influenced only little by the non-linear processes of galaxy formation. This allows us to compute them very precisely (to about 1%, which is high precision for present cosmological standards). For given initial fluctuations, the result depends only on the cosmological parameters. If we can measure CMB anisotropies to a precision of, say 1%, this allows us therefore to determine cosmological parameters to about 1%. An unprecedented possibility! Consider that at present, after the work of two generations, *e.g.* the Hubble parameter is known only to about 25%, the baryon density is known within about a factor of 2 and the uncertainties in the dark matter density, the cosmological constant and the space curvature are even larger.

This somewhat too optimistic conclusion has however three caveats which we want to mention before entering the subject of these lectures.

1. **Initial conditions:** The result depends on the model for the initial fluctuations. Inflationary scenarios contain in general three to four free parameters, like the ratio of tensor to scalar perturbations (r) and the spectral index of the scalar and tensor perturbations (n_S and n_T), so a few more parameters need to be fitted additionally to the data.

   If the perturbations are generated by active sources like, *e.g.*, topological defects, then the modeling is far more complicated, and the analysis is too different to be included in these lectures.

2. **Degeneracy:** Even though we can measure over 1000 independent modes (C_ℓ's) of the CMB anisotropy spectrum, there are certain combinations of the cosmological parameters that lead to degeneracies in the CMB spectrum. The result is, *e.g.*, very sensitive to the sum $\Omega_{\rm matter} + \Omega_\Lambda$, but not to the difference ("cosmic confusion").

3. **Cosmic variance:** Since the fluctuations are created by random processes, we can only calculate expectation values. Yet we have only one universe to take measurements ("cosmic variance"). For small–scale fluctuations we can in general assume that the expectation value over ensembles of universes is the same as a spatial average (a kind of ergodic hypothesis), but for large scales we can't escape large statistical errors.

6.1 Friedmann-Lemaître universes

Friedmann-Lemaître universes are homogeneous and isotropic solutions of Einstein's equations. The hyper-surfaces of constant time are homogeneous and isotropic, *i.e.*, spaces of constant curvature with metric $a^2(\eta)\gamma_{ij}dx^i dx^j$, where γ_{ij} is the metric of a space with constant curvature

κ,

$$\gamma_{ij}dx^i dx^j = dr^2 + \chi^2(r)\left(d\vartheta^2 + sin^2\vartheta d\varphi^2\right) \tag{6.1}$$

$$\chi^2(r) = \begin{cases} r^2 & , \quad \kappa = 0 \\ \sin^2 r & , \quad \kappa = 1 \\ \sinh^2 r & , \quad \kappa = -1, \end{cases} \tag{6.2}$$

where we have rescaled $a(\eta)$ such that $\kappa = \pm 1$ or 0. (With this normalization the scale factor a has the dimension of a length and η and r are dimensionless for $\kappa \neq 0$.) The four-dimensional metric is then of the form

$$g_{\mu\nu}dx^\mu dx^\nu = -a^2(\eta)d\eta^2 + a^2(\eta)\gamma_{ij}dx^i dx^j. \tag{6.3}$$

Here η is called the *conformal time.*

Einstein's equations reduce to ordinary differential equations for the function $a(\eta)$ (with $\dot{} \equiv d/d\eta$):

$$\left(\frac{\dot a}{a}\right)^2 + \kappa = \frac{8\pi G}{3}a^2\rho + \frac{1}{3}\Lambda a^2 \tag{6.4}$$

$$\left(\frac{\dot a}{a}\right)^{\cdot} = -4\pi G a^2(\rho + 3p) + \frac{1}{3}\Lambda a^2 = \left(\frac{\ddot a}{a}\right) - \left(\frac{\dot a}{a}\right)^2, \tag{6.5}$$

where $\rho = -T^0_0$, $p = T^i_i$ (no sum!) and all other components of the energy momentum tensor have to vanish by the requirement of isotropy and homogeneity. Λ is the cosmological constant.

Energy momentum "conservation" (which is also a consequence of (6.4) and (6.5) due to the contracted Bianchi identity) reads

$$\dot\rho = -3\left(\frac{\dot a}{a}\right)(\rho + p). \tag{6.6}$$

After these preliminaries (which we suppose to be known to the audience) let us answer the following question: Given an object with comoving diameter λ[1] at a redshift $z(\eta) = (a_0/a) - 1$. Under which angle $\vartheta(\lambda, z)$ do we see this object today and how does this angle depend on Ω_Λ and Ω_κ?

We define

$$\begin{aligned} \Omega_m &= \left(\frac{8\pi G\rho a^2}{3\left(\frac{\dot a}{a}\right)^2}\right)_{\eta=\eta_0} \\ \Omega_\Lambda &= \left.\frac{\Lambda a^2}{3\left(\frac{\dot a}{a}\right)^2}\right|_{\eta=\eta_0} \\ \Omega_\kappa &= \left.\frac{-\kappa}{\left(\frac{\dot a}{a}\right)^2}\right|_{\eta=\eta_0}, \end{aligned} \tag{6.7}$$

where the index $_0$ indicates the value of a given variable today. Friedmann's equation (6.4) then requires

$$1 = \Omega_m + \Omega_\Lambda + \Omega_\kappa. \tag{6.8}$$

Back to our problem: Without loss of generality we set $r = 0$ at our position and thus $r = r_1 = \eta_0 - \eta_1$ at the position of the flashes, A and B at redshift z_1. If λ denotes the comoving arc length between A and B we have $\lambda = \chi(r_1)\vartheta = \chi(\eta_0 - \eta_1)\vartheta$, *i.e.*

$$\vartheta = \frac{\lambda}{\chi(\eta_0 - \eta_1)}. \tag{6.9}$$

[1] or physical size $a(\eta)\lambda = d$

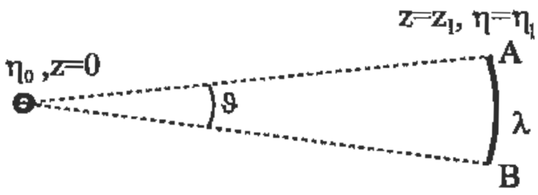

Figure 6.1: The two ends of the object emit a flash simultaneously from A and B at z_1 which reaches us today.

It remains to calculate $(\eta_0 - \eta_1)(z_1)$.

Note that in the case $\kappa = 0$ we can still normalize the scale factor a as we want, and it is convenient to choose $a_0 = 1$, so that comoving scales today become physical scales. However, for $\kappa \neq 0$, we have already normalized a such that $\kappa = \pm 1$ and $\chi = \sin r$ or $\sinh r$. We have in principle no normalisation constant left.

From the Friedmann equation we have

$$\dot{a}^2 = \frac{8\pi G}{3}a^4\rho + \frac{1}{3}\Lambda a^4 - \kappa a^2. \tag{6.10}$$

We assume that ρ is a combination of "dust" (cold, non–relativistic matter) with $p_d = 0$ and radiation with $p_{\rm rad} = {}^1\!/_3\rho_{\rm rad}$.

From (6.6) we find that $\rho_{\rm rad} \propto a^{-4}$ and $\rho_d \propto a^{-3}$. Therefore, with $H_0 = \left(\frac{\dot{a}}{a^2}\right)(\eta_0)$, we define

$$\frac{8\pi G}{3}a^4\rho = H_0^2\left(a_0^4\Omega_{\rm rad} + \Omega_d a a_0^3\right) \tag{6.11}$$

$$\frac{1}{3}\Lambda a^4 = H_0^2\Omega_\Lambda a^4 \tag{6.12}$$

$$-\kappa a^2 = H_0^2\Omega_\kappa a^2 a_0^2\ . \tag{6.13}$$

The Friedmann equation then implies

$$\frac{da}{d\eta} = H_0 a_0^2\left(\Omega_{\rm rad} + \frac{a}{a_0}\Omega_d + \frac{a^4}{a_0^4}\Omega_\Lambda + \frac{a^2}{a_0^2}\Omega_\kappa\right)^{\frac{1}{2}} \tag{6.14}$$

so that

$$\eta_0 - \eta_1 = \frac{1}{H_0 a_0}\int_0^{z_1}\frac{dz}{[\Omega_{\rm rad}(z+1)^4 + \Omega_d(z+1)^3 + \Omega_\Lambda + \Omega_\kappa(z+1)^2]^{\frac{1}{2}}}\ . \tag{6.15}$$

Here we have introduced the cosmological redshift $z + 1 = a_0/a$. (In principle we could of course also add other matter components like, *e.g.* "quintessence" [58], which would lead to a somewhat different form of the integral (6.15), but for definiteness, we remain with dust, radiation and a cosmological constant.)

In general, this integral has to be solved numerically. It determines the angle $\vartheta(\lambda, z_1)$ under which an object with comoving size λ at z_1 is seen.

On the other hand, the angular diameter distance to an object of physical size d seen under angle ϑ is given by $\eta_0 - \eta_1 = r_1 = \chi^{-1}(d/a_1/\vartheta)$. If we are able to measure the redshift and the angular diameter distance of a certain class of objects comparing with Eq. (6.15) allows in principle to determine the parameters Ω_m, Ω_Λ, Ω_κ and H_0.

We have $\frac{-\kappa}{H_0^2 a_0^2} = \Omega_\kappa \Rightarrow H_0 a_0 = \frac{1}{\sqrt{|\Omega_\kappa|}}$ for $\Omega_\kappa \neq 0$.

Observationally we know $10^{-5} < \Omega_{\rm rad} \leq 10^{-4}$ as well as $0.1 \leq \Omega_d \lesssim 1$, $|\Omega_\Lambda| \lesssim 1$ and $|\Omega_\kappa| \lesssim 1$.

If we are interested in small redshifts, $z_1 \lesssim 10$, we may safely neglect $\Omega_{\rm rad}$. In this region, Eq. (6.15) is very sensitive to Ω_Λ and provides an excellent mean to constrain the cosmological constant (see part I).

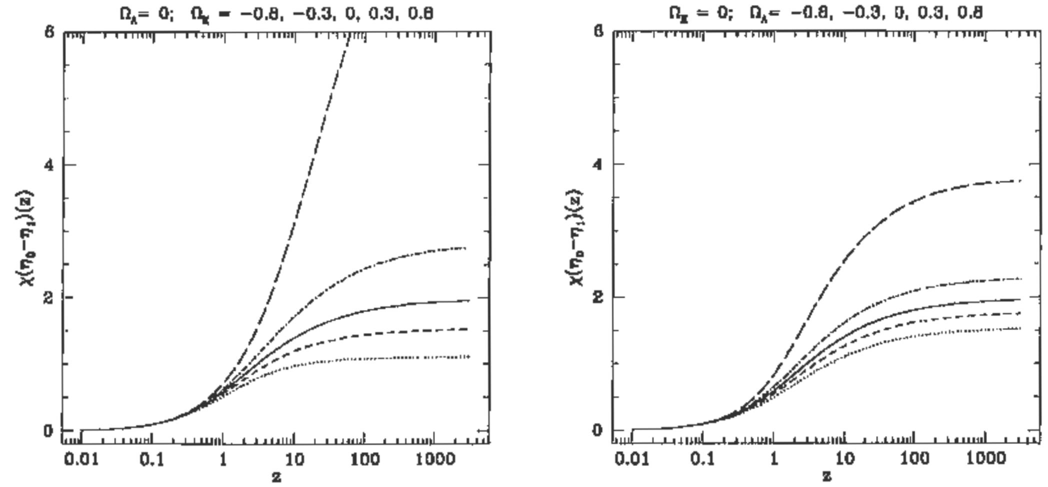

Figure 6.2: The function $\chi(\eta_0 - \eta_1)$ as a function of the redshift z for different values of the cosmological parameters Ω_κ (left, with Ω_Λ=0) and Ω_Λ (right, with Ω_κ=0), namely −0.8 [dotted], −0.3 [short–dashed], 0 [solid], 0.3 [dot–dashed], 0.8 [long–dashed].

At high redshift, $z_1 \gtrsim 1000$, neglecting radiation is no longer a good approximation.

We shall later need the opening angle of the *horizon* distance,

$$\vartheta_H(z_1) = \frac{\eta_1}{\chi(\eta_0 - \eta_1)}, \tag{6.16}$$

$$\eta_1 = \frac{1}{H_0 a_0} \int_{z_1}^{\infty} \frac{dz}{[\Omega_{\rm rad}(z+1)^4 + \Omega_d(z+1)^3 + \Omega_\Lambda + \Omega_\kappa(z+1)^2]^{\frac{1}{2}}}. \tag{6.17}$$

(Clearly this integral diverges if $\Omega_{\rm rad} = \Omega_d = 0$. This is exactly what happens during an inflationary period and leads there to the solution of the horizon problem.)

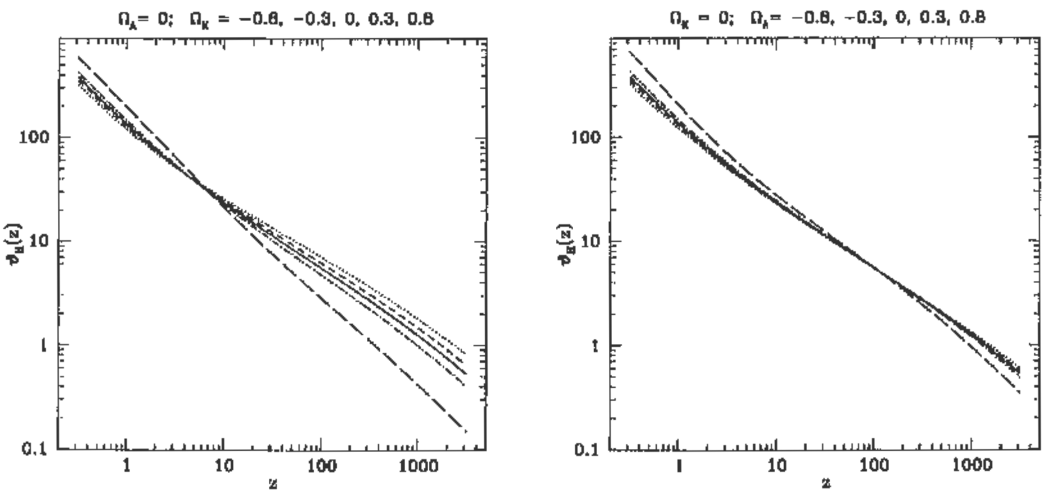

Figure 6.3: $\vartheta_H(z_1)$ (in degrees) for different values of the cosmological parameters Ω_κ and Ω_Λ as in Fig. 6.2.

The value of the radiation density is well known,

$$\rho_{\rm rad} = 7.94 \times 10^{-34}(T_0/2.737{\rm K})^4{\rm g/cm}^3 \ .$$

This gives

$$\Omega_{\rm rad}h^2 = 4.2 \cdot 10^{-5}(T_0/2.737{\rm K})^4 \ , \tag{6.18}$$

$$H_0 = 100h\frac{\rm km}{\rm sMpc} \ . \tag{6.19}$$

Exercise: Neglecting $\Omega_{\rm rad}$, show that for $\Omega_\Lambda = 0$ and small curvature, $0 < |\Omega_\kappa| \ll \Omega_d$ at high enough redshift, $z_1 \geq 10$, $\eta_0 - \eta_1 \simeq 2\sqrt{|\Omega_\kappa|/\Omega_d}$. Conclude that $\vartheta(\lambda, z_1) \propto \sqrt{\Omega_d}$ so that $\vartheta \simeq \sqrt{\Omega_d}\vartheta_1(\lambda, z)$. Calculate $\vartheta_1(\lambda, z) = \vartheta(\lambda, z)|_{\Omega_d=1,\Omega_\kappa=\Omega_\Lambda=\Omega_{\rm rad}=0}$ explixitely.

6.2 Recombination and the cosmic microwave background (CMB)

During its expansion, the universe cools adiabatically. At early times, it is dominated by a thermal radiation background with $\rho = C/a^4 = g_{\rm eff}a_{\rm SB}T^4$,[2] and we find that $T \propto a^{-1}$. Here $g_{eff} = n_b + 7/8n_F$ is the effective number of degrees of freedom, bosons counting as 1 and fermions counting as 7/8 (see *e.g.* [59]). At temperatures below 0.5MeV only neutrinos and photons are still relativistic leading to the density parameter given in Eq. (6.18). (Neutrinos have a somewhat lower temperature than photons, $T_\nu = (4/11)^{1/3}T$, since they have already dropped out of thermal equilibrium before $e^\pm$ annihilation which therefore reheats the photons but not the neutrinos, see *e.g.* [59, 60].)

The photons obey a Planck distribution,

$$f(\omega) = \frac{1}{e^{\omega/T} - 1}. \tag{6.20}$$

At a temperature of about $T \sim 4000{\rm K} \sim 0.4{\rm eV}$, the number density of photons with energies above the hydrogen ionisation energy drops below the baryon density of the universe, and the protons begin to (re-)combine to neutral hydrogen. (Helium has already recombined earlier.) Photons and baryons are tightly coupled before (re-)combination by non-relativistic Thomson scattering of electrons. During recombination the free electron density drops sharply and the mean free path of the photon grows larger than the Hubble scale. At the temperature $T_{\rm dec} \sim 3000{\rm K}$ (corresponding to the redshift $z_{\rm dec} \simeq 1100$ and the physical time $t_{\rm dec} = a_0\eta_{\rm dec} \simeq 10^5$years) photons become *free* and the universe becomes transparent.

After recombination, the photon distribution evolves according to Liouville's equation (geodesic spray):

$$p^0\partial_\eta f - \Gamma^i_{\mu\nu}p^\mu p^\nu\frac{\partial f}{\partial p^i} \equiv L_{X_g}f = 0, \tag{6.21}$$

where $i = 1, 2, 3$. Since the photons are massless, $|\mathbf{p}|^2 = \sum_{i=1}^3 p_ip^i = \omega^2 \ \ (\omega = ap^0)$. Isotropy of the distribution implies that f depends on p^i only via $|\mathbf{p}| = \omega$, and so

$$\frac{\partial f}{\partial p^i} = \frac{\partial\omega}{\partial p^i}\frac{\partial f}{\partial\omega} = \frac{p^i}{\omega}\frac{\partial f}{\partial\omega}. \tag{6.22}$$

In a Friedmann universe (also if $\kappa \neq 0$!) we find for $p^\mu p_\mu = -\omega^2 + \mathbf{p}^2 = 0$ [exercise!]

$$\Gamma^i_{\mu\nu}p^\mu p^\nu p_i = -\omega^3\left(\frac{\dot a}{a^2}\right). \tag{6.23}$$

[2] We will use units with $\hbar = c = k_{\rm B} = 1$ throughout these lectures. The Stefan–Boltzmann constant is then given by $a_{\rm SB} = \pi^2k_{\rm B}^4/(60\hbar^3c^2) = \pi^2/60$.

Inserting this result into (6.21) leads to

$$\partial_\eta f + \omega \left(\frac{\dot{a}}{a} \right) \frac{\partial f}{\partial \omega} = 0, \tag{6.24}$$

which is satisfied by an arbitrary function $f = f(\omega a)$. Hence the distribution of free–streaming photons changes just by redshifting the momenta. Therefore, setting $T \propto a^{-1}$ even after recombination, the blackbody shape of the photon distribution remains unchanged.

Note however that after recombination the photons are no longer in thermal equilibrium and the T in the Planck distribution is not a temperature in the thermodynamical sense but merely a parameter in the photon distribution function.

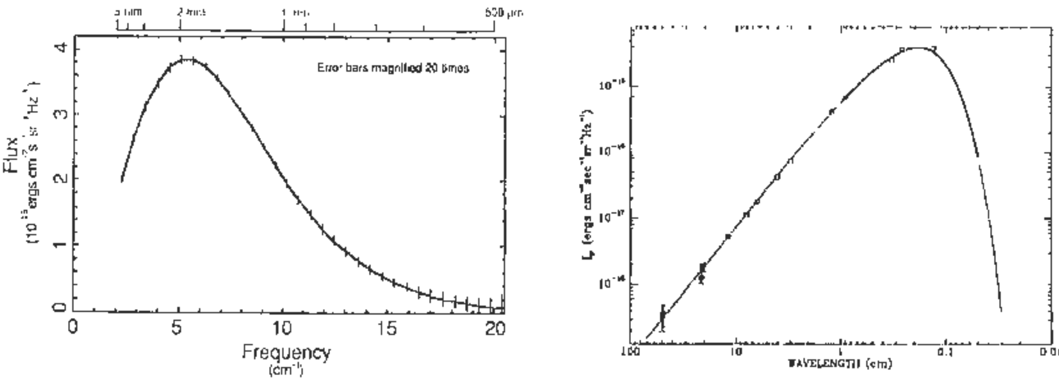

Figure 6.4: Spectrum of the cosmic background radiation. The graph on the left shows the measurements of the FIRAS experiment on COBE (the vertical bars), overlaid by a blackbody spectrum at a temperature of 2.73 K. The error bars are 20 times magnified! The image on the right shows a larger number of measurements. The FIRAS data is represented by the fat line around the peak of the spectrum [60].

The blackbody spectrum of these cosmic photons which are called the "cosmic microwave background" (CMB) is extremely well verified observationally (see Fig. 6.4). The limits on deviations are often parameterized in terms of three parameters: The chemical potential μ, the Compton y parameter (which quantifies a well defined change in the spectrum arising from interactions with a non–relativistic electron gas at a different temperature, see *e.g.* [60]) and Y_{ff} (describing a contamination by free-free emission).

The present limits on these parameters are (at 95% CL, [56])

$$|\mu| < 9 \cdot 10^{-5}, \quad |y| < 1.2 \cdot 10^{-5}, \quad |Y_{ff}| < 1.9 \cdot 10^{-5}. \tag{6.25}$$

The CMB Photons have not only a very thermal spectrum, but they are also distributed very isotropically, apart from a dipole which is (most probably) simply due to our motion relative to the surface of last scattering:

An observer moving with velocity $\mathbf{v}$ relative to a source emitting a photon with proper momentum $\mathrm{p} = -\omega \mathrm{n}$ sees this photon redshifted with frequency

$$\omega' = \gamma\omega \left(1 - \mathbf{n}\mathbf{v}\right), \tag{6.26}$$

in first order in $\mathbf{v}$ this is just a dipole perturbation. This dipole anisotropy, which is of the order of

$$\left(\frac{\Delta T}{T} \right)_{\text{dipole}} \simeq 10^{-3}$$

has already been discovered in the 70ties [61, 62]. Interpreting it as due to our motion with respect to the last scattering surface implies a velocity for the solar-system bary-center of $v = 371 \pm 0.5\,\mathrm{km/s}$ at 68% CL ([56]).

The COBE[3] DMR experiment (Differential Microwave Radiometer) has found fluctuations of

$$\sqrt{\left\langle \left(\frac{\Delta T}{T}\right)^2 \right\rangle} \sim 10^{-5} \tag{6.27}$$

on all angular scales $\theta \geq 7°$ [57]. On smaller angular scales many experiments have found fluctuations (that we shall describe in detail later), but all of them are $\lesssim 10^{-4}$.

As we shall see later, the CMB fluctuations on large scales provide a measure for the deviation of the geometry from the Friedmann-Lemaître one. The geometry perturbations are thus small and we may calculate their effects by ***linear perturbation theory.*** On smaller scales, $\Delta T/T$ reflects the fluctuations in the energy density in the baryon/radiation plasma prior to recombination. Their amplitude is just about right to allow the formation of the presently observed non–linear structures (like galaxies, clusters, etc.) out of small initial fluctuations by gravitational instability.

These findings strongly support the hypothesis which we will assume during these lectures, namely that the large scale structure (*i.e.* galaxy distribution) observed in the universe formed by gravitational instability from relatively small ($\sim 10^{-4} - 10^{-5}$) initial fluctuations. As we shall see, such initial fluctuations leave an interesting "fingerprint" on the cosmic microwave background.

[3]Cosmic Background Explorer, NASA satellite launched 1990.

Chapter 7

Perturbation Theory

The tool for the analysis of CMB anisotropies is cosmological perturbation theory. We spend therefore some time on this subject, especially on the fundamental level.

Once all the variables are defined, we will be rather brief in what concerns the derivation of the basic perturbation equations. First of all, because these derivations are in general not very illuminating and secondly because nowadays all of you can obtain them very easily by setting

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \varepsilon a^2 h_{\mu\nu} \tag{7.1}$$

($\bar{g}_{\mu\nu}$ being the unperturbed Friedmann metric) and asking Mathematic or Maple to calculate the Einstein Tensor using the condition $\varepsilon^2 = 0$. We conventionally set (absorbing the "smallness" parameter ε into $h_{\mu\nu}$)

$$\begin{array}{llll} g_{\mu\nu} = \bar{g}_{\mu\nu} + a^2 h_{\mu\nu}, & \bar{g}_{00} = -a^2, & \bar{g}_{ij} = a^2\gamma_{ij} & |h_{\mu\nu}| \ll 1 \\ T^\mu_\nu = \overline{T}^\mu_\nu + \theta^\mu_\nu, & \overline{T}^0_0 = -\bar{\rho}, & \overline{T}^i_j = \bar{p}\delta^i_j & |\theta^\mu_\nu|/\bar{\rho} \ll 1. \end{array} \tag{7.2}$$

7.1 Gauge transformation, gauge invariance

The first fundamental problem we want to discuss is the problem of 'choice of gauge' in cosmological perturbation theory:

For linear perturbation theory to apply, the spacetime manifold $\mathcal{M}$ with metric g and the energy momentum tensor T of the real, observable universe must be in some sense close to a Friedmann universe, *i.e.*, the manifold $\mathcal{M}$ with a Robertson–Walker metric $\bar{g}$ and a homogeneous and isotropic energy momentum tensor $\overline{T}$. It is an interesting, non–trivial unsolved problem how to construct $\bar{g}$ and $\overline{T}$ from the physical fields g and T in practice. There are two main difficulties: Spatial averaging procedures depend on the choice of a hyper–surface of constant time and do not commute with derivatives, so that averaged fields $\bar{g}$ and $\overline{T}$ will in general not satisfy Einstein's equations. Secondly, averaging is in practice impossible over super–horizon scales.

Even though we cannot give a constructive prescription, we now assume that there exists an averaging procedure which leads to a Friedmann universe with spatially averaged tensor fields $\overline{Q}$, such that the deviations $(T_{\mu\nu} - \overline{T}_{\mu\nu})/\max_{\{\alpha\beta\}}\{|\overline{T}_{\alpha\beta}|\}$ and $(g_{\mu\nu} - \overline{g}_{\mu\nu})/\max_{\{\alpha\beta\}}\{\overline{g}_{\alpha\beta}\}$ are small, and $\bar{g}$ and $\overline{T}$ satisfy Friedmann's equations. Let us call such an averaging procedure 'admissible'. There may be many other admissible averaging procedures (e.g. over a different hyper–surface) leading to slightly differeut Friedmann backgrounds. But since $|g - g|$ is small of order ϵ, the difference of the two Friedmann backgrounds must also be small of order ϵ and we can regard it as part of the perturbation.

We consider now a fixed admissible Friedmann background $(\bar{g}, \overline{T})$ as chosen. Since the theory is invariant under diffeomorphisms (coordinate transformations), the perturbations are not unique. For an arbitrary diffeomorphism ϕ and its pullback ϕ^*, the two metrics g and $\phi^*(g)$ describe the same geometry. Since we have chosen the background metric $\bar{g}$ we only allow diffeomorphisms which leave $\bar{g}$ invariant *i.e.* which deviate only in first order form the identity. Such an 'infinitesimal' isomorphism can be represented as the infinitesimal flow of a vector field X, $\phi = \phi^X_\epsilon$. Remember the definition of the flow: For the integral curve $\gamma_x(s)$ of X with starting point x, i.e., $\gamma_x(s=0) = x$ we have $\phi^X_s(x) = \gamma_x(s)$. In terms of the vector field X, to first order in ϵ, its pullback is then of the form

$$\phi^* = id + \epsilon L_X$$

(L_X denotes the Lie derivative in direction X). The transformation $g \to \phi^*(g)$ is equivalent to $\bar{g} + \epsilon a^2 h \to \bar{g} + \epsilon(a^2 h + L_X \bar{g})$, *i.e.* under an 'infinitesimal coordinate transformation' the metric perturbation h transforms as

$$h \to h + a^{-2} L_X \bar{g} \ . \tag{7.3}$$

In the context of cosmological perturbation theory, infinitesimal coordinate transformations are called 'gauge transformation'. The perturbation of a arbitrary tensor field $Q = \bar{Q} + \epsilon Q^{(1)}$ obeys the gauge transformation law

$$Q^{(1)} \to Q^{(1)} + L_X \bar{Q} \ . \tag{7.4}$$

Since every vector field X generates a gauge transformation $\phi = \phi^X_\epsilon$, we can conclude that only perturbations of tensor fields with $L_X \overline{Q} = 0$ for all vector fields X, i.e., with vanishing (or constant) 'background contribution' are gauge invariant. This simple result is sometimes referred to as the 'Stewart Walker Lemma' [52].

The gauge dependence of perturbations has caused many controversies in the literature, since it is often difficult to extract the physical meaning of gauge dependent perturbations, especially on super–horizon scales. This has led to the development of gauge invariant perturbation theory which we are going to use throughout these lectures. The advantage of the gauge–invariant formalism is that the variables used have simple geometric and physical meanings and are not plagued by gauge modes. Although the derivation requires somewhat more work, the final system of perturbation equations is usually simple and well suited for numerical treatment. We shall also see, that on sub-horizon scales, the gauge invariant matter perturbations variables approach the usual, gauge dependent ones, and one of the geometrical variables corresponds to the Newtonian potential, so that the Newtonian limit can be performed easily.

First we note that since all relativistic equations are covariant (i.e. can be written in the form $Q = 0$ for some tensor field Q), it is always possible to express the corresponding perturbation equations in terms of gauge invariant variables [53, 54, 55].

7.2 Gauge invariant perturbation variables

Since the $\{\eta = \text{const}\}$ hyper-surfaces are homogeneous and isotropic, it is sensible to perform a harmonic analysis: A (spatial) tensor field Q on these hyper-surfaces can be decomposed into components with transform irreducibly under translations and rotations. All such components evolve independently. For a scalar quantity f in the case $\kappa = 0$ this is nothing else than its Fourier decomposition:

$$f(\mathbf{x}, \eta) = \int d^3k \hat{f}(\mathbf{k}) e^{i\mathbf{kx}}. \tag{7.5}$$

(The exponentials $Y_\mathbf{k}(\mathbf{x}) = e^{i\mathbf{kx}}$ are the unitary irreducible representations of the Euclidean translation group.) For $\kappa = 1$ such a decomposition also exists, but the values k are discrete, $k^2 = \ell(\ell+2)$ and for $\kappa = -1$, they are bounded from below, $k^2 > 1$. Of course, the functions $Y_\mathbf{k}$ are different for $\kappa \neq 0$.

They are always the complete orthogonal set of eigenfunctions of the Laplacian,

$$\Delta Y^{(S)} = -k^2 Y^{(S)}. \tag{7.6}$$

In addition, a variable (at fixed position x) can be decomposed into irreducible components under the rotation group $SO(3)$.

For a vector field, this is its decomposition into a gradient and a rotation,

$$V_i = \nabla_i \varphi + B_i, \tag{7.7}$$

where

$$B^i_{|i} = 0, \tag{7.8}$$

where we used $X_{|i}$ to denote the three-dimensional covariant derivative of X. φ is the spin 0 and **B** is the spin 1 component of V.

For a symmetric tensor field we have

$$H_{ij} = H_L\gamma_{ij} + \left(\nabla_i\nabla_j - \frac{1}{3}\Delta\gamma_{ij}\right)H_T + \frac{1}{2}\left(H^{(V)}_{i|j} + H^{(V)}_{j|i}\right) + H^{(T)}_{ij}, \tag{7.9}$$

where

$$H^{(V)|i}_i = H^{(T)^i}_i = H^{(T)^j}_{i|j} = 0. \tag{7.10}$$

Here H_L and H_T are spin 0 components, $H^{(V)}_i$ is a spin 1 component and $H^{(T)}_{ij}$ is a spin 2 component.

We shall not need higher tensors (or spinors) in these lectures. As a basis for vector and tensor modes we use the vector and tensor type eigenfunctions to the Laplacian,

$$\Delta Y^{(V)}_j = -k^2 Y^{(V)}_j \tag{7.11}$$

and

$$\Delta Y^{(T)}_{ji} = -k^2 Y^{(T)}_{ji}\ , \tag{7.12}$$

where $Y^{(V)}_j$ is a transeverse vector, $Y^{(V)|j}_j = 0$ and $Y^{(T)}_{ji}$ is a symetric transverse traccless tensor, $Y^{(T)j}_j = Y^{(T)|i}_{ji} = 0$.

According to Eqs. (7.7,7.9) we can construct scalar type vectors and tensors and vector type tensors. To this goal we define

$$Y^{(S)}_j \equiv -k^{-1}Y^{(S)}_{|j} \tag{7.13}$$

$$Y^{(S)}_{ij} \equiv k^{-2}Y^{(S)}_{|ij} + \frac{1}{3}\gamma_{ij}Y^{(S)} \tag{7.14}$$

$$Y^{(V)}_{ij} \equiv -\frac{1}{2k}(Y^{(V)}_{i|j} + Y^{(V)}_{j|i})\ . \tag{7.15}$$

In the following we shall extensively use this decomposition and write down the perturbation equations for a given mode k.

The decomposition of a vector field is then of the form

$$B_i = BY^{(S)}_i + B^{(V)}Y^{(V)}_i. \tag{7.16}$$

The decomposition of a scalar field is given by (compare 7.9)

$$H_{ij} = H_L Y^{(S)}\gamma_{ij} + H_T Y^{(S)}_{ij} + H^{(V)}Y^{(V)}_{ij} + H^{(T)}Y^{(T)}_{ij}, \tag{7.17}$$

7.2.1 Metric perturbations

Perturbations of the metric are of the form

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + a^2 h_{\mu\nu}. \tag{7.18}$$

We parameterize them as

$$h_{\mu\nu}dx^\mu dx^\nu = -2Ad\eta^2 + 2B_i d\eta dx^i + 2H_{ij}dx^i dx^j, \tag{7.19}$$

and we decompose the perturbation variables B_i and H_{ij} according to (7.16) and (7.17).

Let us consider the behaviour of $h_{\mu\nu}$ under gauge transformations. We set the vector field defining the gauge transformation to

$$\mathbf{X} = T\partial_\eta + L^i\partial_i. \tag{7.20}$$

Using simple identities from differential geometry like $L_{\mathbf{X}}(df) = d(L_{\mathbf{X}}f)$ and $(L_{\mathbf{X}}\gamma)_{ij} = X_{i|j} + X_{j|i}$, we obtain

$$\begin{aligned} L_{\mathbf{X}}\bar{g} &= a^2\left[-2\left(\frac{\dot{a}}{a}T + \dot{T}\right)d\eta^2 + 2\left(\dot{L}_i - T_{,i}\right)d\eta dx^i \right. \\ &\left. + \left(2\frac{\dot{a}}{a}T\gamma_{ij} + L_{i|j} + L_{j|i}\right)dx^i dx^j\right]. \end{aligned} \tag{7.21}$$

Comparing this with (7.19) and using (7.3) we obtain the following behaviour of our perturbation variables under gauge transformations (decomposing $L_i = LY_i^{(S)} + L^{(V)}Y_i^{(V)}$):

$$A \rightarrow A + \frac{\dot{a}}{a}T + \dot{T} \tag{7.22}$$
$$B \rightarrow B - \dot{L} - kT \tag{7.23}$$
$$B^{(V)} \rightarrow B^{(V)} - \dot{L}^{(V)} \tag{7.24}$$
$$H_L \rightarrow H_L + \frac{\dot{a}}{a}T + \frac{k}{3}L \tag{7.25}$$
$$H_T \rightarrow H_T - kL \tag{7.26}$$
$$H^{(V)} \rightarrow H^{(V)} - kL^{(V)} \tag{7.27}$$
$$H^{(T)} \rightarrow H^{(T)}. \tag{7.28}$$

Two scalar and one vector variable can be brought to disappear by gauge transformations.

One often chooses $kL = H_T$ and $T = B + \dot{L}$, so that the variables H_T and B vanish. In this gauge (longitudinal gauge), scalar perturbations of the metric are of the form ($H_T = B = 0$):

$$h_{\mu\nu}^{(S)} = -2\Psi d\eta^2 + 2\Phi\gamma_{ij}dx^i dx^j. \tag{7.29}$$

Ψ and Φ are the so called *Bardeen* potentials. In general they are given by

$$\Psi = A - \frac{\dot{a}}{a}k^{-1}\sigma - k^{-1}\dot{\sigma} \tag{7.30}$$
$$\Phi = H_L + \frac{1}{3}H_T - \frac{\dot{a}}{a}k^{-1}\sigma \tag{7.31}$$

with $\sigma = k^{-1}\dot{H}_T - B$. A short calculation shows that they are gauge invariant.

For vector perturbations it is convenient to set $kL^{(V)} = H^{(V)}$ so that $H^{(V)}$ vanishes and we have

$$h_{\mu\nu}^{(V)}dx^\mu dx^\nu = -2\sigma^{(V)}Y_i^{(V)}d\eta dx^i. \tag{7.32}$$

We shall call this gauge the "vector gauge". In general $\sigma^{(V)} = k^{-1}\dot{H}^{(V)} - B^{(V)}$ is gauge invariant[1].

Clearly there are no tensorial (spin 2) gauge transformation and hence $H_{ij}^{(T)}$ is gauge invariant.

7.2.2 Perturbations of the energy momentum tensor

Let $T^\mu_\nu = \overline{T}^\mu_\nu + \Theta^\mu_\nu$ be the full energy momentum tensor. We define its energy density ρ and its energy flow 4-vector u as the timelike eigenvalue and eigenvector of T^μ_ν:

$$T^\mu_\nu u^\nu = -\rho u^\mu, \quad u^2 = -1. \tag{7.33}$$

We then define their perturbations by

$$\rho = \bar{\rho}(1+\delta), \quad u = u^0\partial_t + u^i\partial_i. \tag{7.34}$$

[1] $Y_{ij}^{(V)}\sigma^{(V)}$ is the shear of the hyper-surfaces of constant time.

u^0 is fixed by the normalisation condition,

$$u^0 = \frac{1}{a}(1 - A). \tag{7.35}$$

We further set

$$u^i = \frac{1}{a}v_i = vY^i_{(S)} + v^{(V)}Y^i_{(V)}. \tag{7.36}$$

We define $P^\mu_\nu \equiv u^\mu u_\nu + \delta^\mu_\nu$, the projection tensor onto the part of tangent space normal to u and set the stress tensor

$$\tau^{\mu\nu} = P^\mu_\alpha P^\nu_\beta T^{\alpha\beta}. \tag{7.37}$$

In the unperturbed case we have $\tau^0_0 = 0, \tau^i_j = \bar{p}\delta^i_j$. Including perturbations, to first order we still obtain

$$\tau^0_0 = \tau^0_i = \tau^i_0 = 0. \tag{7.38}$$

But τ^i_j contains in general perturbations. We set

$$\tau^i_j = \bar{p}\left[(1 + \Pi_L)\,\delta^i_j + \Pi^i_j\right], \quad \text{with} \quad \Pi^i_i = 0. \tag{7.39}$$

We decompose Π^i_j as

$$\Pi^i_j = \Pi^{(S)}Y^{(S)\,i}_j + \Pi^{(V)}Y^{(V)\,i}_j + \Pi^{(T)}Y^{(T)\,i}_j. \tag{7.40}$$

We shall not go in detail through the gauge transformation properties, but just state some results which can be obtained as an exercise:

- Of the variables defined above only the $\Pi^{(S,V,T)}$ are gauge invariant; they describe the anisotropic stress tensor, $\Pi^\mu_\nu = \tau^\mu_\nu - {}^1\!/_3\tau^\alpha_\alpha\delta^\mu_\nu$. They are therefore gauge invariant due to the Stewart–Walker lemma, since $\bar{\Pi} = 0$. For perfect fluids $\Pi^\mu_\nu = 0$.

- A second gauge invariant variable is

$$\Gamma = \pi_L - \frac{c_s^2}{w}\delta, \tag{7.41}$$

  where $c_s^2 \equiv \dot{p}/\dot{\rho}$ is the adiabatic sound speed and $w \equiv p/\rho$ is the enthalpy. One can show that Γ is proportional to the divergence of the entropy flux of the perturbations. Adiabatic perturbations are therefore characterised by $\Gamma = 0$.

- Gauge invariant density and velocity perturbations can be found by combining δ, v and $v_i^{(V)}$ with metric perturbations.

We shall use

$$V \equiv v - \frac{1}{k}\dot{H}_T = v^{(\mathrm{long})} \tag{7.42}$$

$$D_g \equiv \delta + 3(1+w)\left(H_L + \frac{1}{3}H_T\right) = \delta^{(\mathrm{long})} + 3(1+w)\Phi \tag{7.43}$$

$$D \equiv \delta^{(\mathrm{long})} + 3(1+w)\left(\frac{\dot{a}}{a}\right)\frac{V}{k} \tag{7.44}$$

$$V^{(V)} \equiv v^{(V)} - \frac{1}{k}\dot{H}^{(V)} = v^{(\mathrm{vec})} \tag{7.45}$$

$$\Omega \equiv v^{(V)} - B^{(V)} = v^{(\mathrm{vec})} - B^{(V)} \tag{7.46}$$

$$\Omega - V^{(V)} = \sigma^{(V)}. \tag{7.47}$$

Here $v^{(\mathrm{long})}$, $\delta^{(\mathrm{long})}$ and $v_i^{(\mathrm{vec})}$ are the velocity (and density) perturbations in the longitudinal and vector gauge respectively and $\sigma^{(V)}$ is the metric perturbation in vector gauge (see Eq. (7.32)).

These variables can be interpreted nicely in terms of gradients of the energy density and the shear and vorticity of the velocity field [63].

But we just want to show that on scales much smaller than the Hubble scale, $k\eta \gg 1$, the metric perturbations are much smaller than δ and v and we can thus "forget them" (which will be important when comparing experimental results with calculations in this formalism):

The perturbations of the Einstein tensor are given by second derivatives of the metric perturbations. Einstein's equations yield the following order of magnitude estimate:

$$\mathcal{O}\left(\frac{\delta T}{T}\right) \underbrace{\mathcal{O}\left(8\pi G T\right)}_{\mathcal{O}\left(\frac{\dot a}{a}\right)^2 = \mathcal{O}(\eta^{-2})} = \mathcal{O}\left(\frac{1}{\eta^2}h + \frac{k}{\eta}h + k^2 h\right) \tag{7.48}$$

$$\mathcal{O}\left(\frac{\delta T}{T}\right) = \mathcal{O}\left(h + k\eta h + (k\eta)^2 h\right). \tag{7.49}$$

For $k\eta \gg 1$ this gives $\mathcal{O}(\delta, v) = \mathcal{O}\left(\frac{\delta T}{T}\right) \gg \mathcal{O}(h)$. On sub-horizon scales the difference between δ, $\delta^{(\mathrm{long})}$, D_g and D is negligible as well as the difference between v and V or $v^{(V)}$, $V^{(V)}$ and $\Omega^{(V)}$.

Later we shall also need other perturbation variables like the perturbation of the photon brightness (energy–integrated photon distribution function), but we shall introduce them as we get there and discuss some applications first.

7.3 Basic perturbation equations

As already announced, we do not derive Einstein's equations but just write down those which we shall need later:

7.3.1 Constraint equations

$$\left.\begin{array}{rcll} 4\pi G a^2 \rho D &=& (k^2 - 2\kappa)\Phi & (00) \\ 4\pi G a^2 (\rho + p) V &=& k\left(\left(\frac{\dot a}{a}\right)\Psi - \dot\Phi\right) & (0i) \end{array}\right\} \quad \text{(scalar)} \tag{7.50}$$

$$8\pi G a^2 (\rho + p)\Omega = \frac{1}{2}\left(2\kappa - k^2\right)\sigma^{(V)} \quad (0i) \qquad \text{(vector)} \tag{7.51}$$

7.3.2 Dynamical equations

$$-k^2\left(\Phi + \Psi\right) = 8\pi G a^2 p \Pi^{(S)} \qquad \text{(scalar)} \tag{7.52}$$

$$k\left(\dot\sigma^{(V)} + 2\left(\frac{\dot a}{a}\right)\sigma^{(V)}\right) = 8\pi G a^2 p \Pi^{(V)} \qquad \text{(vector)} \tag{7.53}$$

$$\ddot H^{(T)} + 2\left(\frac{\dot a}{a}\right)\dot H^{(T)} + \left(2\kappa + k^2\right) H^{(T)} = 8\pi G a^2 p \Pi^{(T)}_{ij} \qquad \text{(tensor)} \tag{7.54}$$

Note that for perfect fluids, where $\Pi^i_j \equiv 0$, we have $\Phi = -\Psi$, $\sigma^{(V)} \propto 1/a^2$ and H obeys a damped wave equation. The damping term can be neglected on small scales (over short time periods) when $\eta^{-2} \lesssim 2\kappa + k^2$, and H_{ij} represents propagating gravitational waves. For vanishing curvature, these are just the sub-horizon scales, $k\eta \gtrsim 1$. For $\kappa < 0$, waves oscillate with a somewhat smaller frequency, $\omega = \sqrt{2\kappa + k^2}$, while for $\kappa > 0$ the frequency is somewhat larger.

7.3.3 Conservation equations

$$\left.\begin{array}{c} \dot D_g + 3\left(c_s^2 - w\right)\left(\frac{\dot a}{a}\right) D_g + (1+w)kV + 3w\left(\frac{\dot a}{\rho}\right)\Gamma = 0 \\ \dot V + \left(\frac{\dot a}{a}\right)\left(1 - 3c_s^2\right) V = k\left(\Psi - 3c_s^2\Phi\right) + \frac{c_s^2 k}{1+w} D_g \\ + \frac{wk}{1+w}\left[\Gamma - \frac{2}{3}\left(1 - \frac{3\kappa}{k^2}\right)\Pi\right] \end{array}\right\} \quad \text{(scalar)} \tag{7.55}$$

$$\dot{\Omega}_i + \left(1 - 3c_s^2\right)\left(\frac{\dot{a}}{a}\right)\Omega_i = \frac{p}{2(\rho + p)}\left(k - \frac{2\kappa}{k}\right)\Pi_i^{(V)} \qquad \text{(vector)} \qquad (7.56)$$

Chapter 8

Simple applications

We first discuss some simple applications which will be important for the CMB. We could of course also write (7.55) in terms of D, but we shall just work with the relation

$$D = D_g + 3(1+w)\left(-\Phi + \left(\frac{\dot{a}}{a}\right)k^{-1}V\right). \tag{8.1}$$

8.1 The pure dust fluid at $\kappa = 0, \Lambda = 0$

We assume the dust to have $w = c_s^2 = p = 0$ and $\Pi = \Gamma = 0$. The equations (7.55), (7.52) and (7.50) then reduce to

$$\dot{D}_g = -kV \quad \text{(energy conservation eqn.)} \tag{8.2}$$

$$\dot{V} + \left(\frac{\dot{a}}{a}\right)V = k\Psi \quad \text{(gravitational acceleration eqn.)} \tag{8.3}$$

$$\Phi = -\Psi \tag{8.4}$$

$$-k^2\Psi = 4\pi G a^2 \rho\left(D_g + 3\left(\Psi + \left(\frac{\dot{a}}{a}\right)k^{-1}V\right)\right) \quad \text{(Poisson eqn.)}. \tag{8.5}$$

In a pure dust universe $\rho \propto a^{-3} \Rightarrow (\dot{a}/a)^2 \propto a^{-1}$, which is solved by $a \propto \eta^2$. The Einstein equations then give immediately $4\pi G\rho a^2 = {}^3\!/_2(\dot{a}/a)^2 = 6/\eta^2$. Setting $k\eta = x$ and $' = d/dx$, the system (8.2-8.5) then becomes

$$D_g' = -V \tag{8.6}$$

$$V' + \frac{2}{x}V = \Psi \tag{8.7}$$

$$\frac{6}{x^2}\left(D_g + 3\left(\Psi + \frac{2}{x}V\right)\right) = -\Psi. \tag{8.8}$$

We use (8.8) to eliminate Ψ and (8.6) to eliminate D_g, leading to

$$\left(18 + x^2\right)V'' + \left(\frac{72}{x} + 4x\right)V' - \left(\frac{72}{x^2} + 4\right)V = 0. \tag{8.9}$$

The general solution is then found to be

$$V = V_0 x + \frac{V_1}{x^4}. \tag{8.10}$$

Since the perturbations are supposed to be small initially, they cannot diverge for $x \to 0$, and we have therefore to choose $V_1 = 0$ (the growing mode). Another way to argue is as follows: If the mode V_1 has to be small already at some early initial time $\eta_{\rm in}$, it will be even much smaller at later times and may hence be neglected. The perturbation variables are then given by

$$V = V_0 x \tag{8.11}$$
$$D_g = -15V_0 - \frac{1}{2}V_0 x^2 \tag{8.12}$$
$$\Psi = 3V_0. \tag{8.13}$$

The constancy of the gravitational potential Ψ in a matter dominated universe and the growth of the density perturbations like the scale factor a led Lifshitz to conclude 1946 [64] that pure gravitational instability cannot be the cause for structure formation: If we start from tiny thermal fluctuations of the order of 10^{-35}, they can only grow to about 10^{-30} through this process during the matter dominated regime. Or, to put it differently, if we do not want to modify the process of structure formation, we need initial fluctuations of the order of 10^{-5}. One possibility to create such fluctuations is due to quantum particle production in the classical gravitational field during inflation. The rapid expansion of the universe blow them up from microscopic scales to cosmological scales.

We distinguish two regimes:
i) super-horizon, $x \ll 1$ where we have

$$D_g = -15V_0 \tag{8.14}$$
$$\Psi = 3V_0 \tag{8.15}$$
$$V = V_0 x \tag{8.16}$$

and *ii)* sub-horizon, $x \gg 1$ where the solution is dominated by the terms

$$V = V_0 x \tag{8.17}$$
$$D_g = -\frac{1}{2}V_0 x^2 \tag{8.18}$$
$$\Psi = 3V_0 = \text{const} \tag{8.19}$$

Exercise: Write and discuss this system in terms of the variables D, V and Ψ. Compare the results!

8.2 The pure radiation fluid, $\kappa = 0, \Lambda = 0$

In this limit we set $w = c_s^2 = 1/3$ and $\Pi = 0$. We conclude from $\rho \propto a^{-4}$ that $a \propto \eta$ and $\Phi = -\Psi$, and the perturbation equations become (with the notation as above):

$$D_g' = -\frac{4}{3}V \tag{8.20}$$
$$V' = 2\Psi + \frac{1}{4}D_g \tag{8.21}$$
$$-2x^2\Psi = 3D_g + 12\Psi + \frac{12}{x}V \tag{8.22}$$

The general solution of this system is

$$\begin{aligned} D_g &= D_2\left[\cos\left(\frac{x}{\sqrt{3}}\right) - 2\frac{\sqrt{3}}{x}\sin\left(\frac{x}{\sqrt{3}}\right)\right] \\ &\quad + D_1\left[\sin\left(\frac{x}{\sqrt{3}}\right) + 2\frac{\sqrt{3}}{x}\cos\left(\frac{x}{\sqrt{3}}\right)\right] \end{aligned} \tag{8.23}$$
$$V = -\frac{3}{4}D_g' \tag{8.24}$$
$$\Psi = \frac{-3D_g - (12/x)V}{12 + 2x^2}. \tag{8.25}$$

Again, regularity at $x = 0$ implies $D_1 = 0$.

In the **super-horizon,** $x \ll 1$ regime we obtain

$$\Psi = \Psi_0, \quad D_g = D_0 - \frac{2}{3}V_0 x^2, \quad V = V_0 x \tag{8.26}$$

with

$$\begin{aligned} D_0 &= -6\Psi_0 = -D_2 & (8.27)\\ V_0 &= \frac{1}{2}\Psi_0 = -\frac{1}{12}D_0. & (8.28)\end{aligned}$$

On **sub-horizon,** $x \gg 1$ scales we find oscillating solutions with constant amplitude with a frequency of $1/\sqrt{3}$:

$$\begin{aligned} V &= V_2 \sin\left(\frac{k\eta}{\sqrt{3}}\right) & (8.29)\\ D_g &= D_2 \cos\left(\frac{k\eta}{\sqrt{3}}\right), \quad \Psi = -\frac{3}{2}x^{-2}D_g & (8.30)\\ D_2 &= \frac{4V_2}{\sqrt{3}}. & (8.31)\end{aligned}$$

We conclude therefore that perturbations outside the Hubble horizon are frozen to first order. Once they enter the horizon they start to collapse, but pressure resists the gravitational force and the radiation fluid starts to oscillate. The perturbations of the gravitational potential oscillate and decay like $1/a^2$ inside the horizon.

8.3 Adiabatic and isocurvature initial conditions for a matter & radiation fluid

In this section we want to investigate a system with a matter and a radiation component that are coupled only through gravity. The matter component acts therefore as *dark matter*, since it does not interact directly with the radiation.

Since the matter and radiation perturbations behave in the same way on super-horizon scales,

$$D_g^{(r)} = A + Bx^2, \quad D_g^{(m)} = A' + B'x^2, \quad V^{(r)} \propto V^{(m)}, \tag{8.32}$$

we may require a constant relation between matter and radiation perturbations. As we have seen in the previous section, inside the horizon ($x > 1$) radiation perturbations start to oscillate while matter perturbations keep following a power law. On sub-horizon scales a constant ratio can thus no longer be maintained. There are two interesting possibilities:

8.3.1 Adiabatic initial conditions

Adiabaticity requires that matter and radiation perturbations are initially in perfect thermal equilibrium. This implies that their velocity fields agree (see below, section of the Boltzmann eqn.!)

$$V^{(r)} = V^{(m)}, \tag{8.33}$$

so the energy flux in the two fluids is coupled initially.

Let us investigate the radiation solution in the *matter dominated era*, when the corresponding scale is already sub-horizon. Since Ψ is dominated by the matter contribution, we have $\Psi \simeq \text{const} = \Psi_0$. We neglect the (decaying) contribution from the sub-dominant radiation to Ψ. Energy-momentum conservation for radiation then gives

$$\begin{aligned} D_g^{(r))\prime} &= -\frac{4}{3}V^{(r)} & (8.34)\\ V^{(r)} &= 2\Psi + \frac{1}{4}D_g^{(r)}. & (8.35)\end{aligned}$$

Now Ψ is just a constant given by the matter perturbations, and we assume that it acts just like a constant source term. The full solution of this system is then found to be

$$D_g^{(r)} = A\cos\left(\frac{x}{\sqrt{3}}\right) - \frac{4}{\sqrt{3}}B\sin\left(\frac{x}{\sqrt{3}}\right) - 8\Psi\left[\cos\left(\frac{x}{\sqrt{3}}\right) - 1\right] \quad (8.36)$$

$$V^{(r)} = B\cos\left(\frac{x}{\sqrt{3}}\right) + \frac{\sqrt{3}}{4}A\sin\left(\frac{x}{\sqrt{3}}\right) - 2\sqrt{3}\Psi\sin\left(\frac{x}{\sqrt{3}}\right). \quad (8.37)$$

Our adiabatic initial conditions require

$$\lim_{x\to 0}\frac{V^{(r)}}{x} = V_0 = \lim_{x\to 0}\frac{V^{(m)}}{x} < \infty. \quad (8.38)$$

Therefore $B = 0$ and $A = 4V_0 - 8\Psi$. Using in addition $\Psi = 3V_0$ (see (8.19)) we obtain

$$D_g^{(r)} = -\frac{44}{3}\Psi\cos\left(\frac{x}{\sqrt{3}}\right) + 8\Psi \quad (8.39)$$

$$V^{(r)} = \frac{1}{\sqrt{3}}\Psi\sin\left(\frac{x}{\sqrt{3}}\right) \quad (8.40)$$

$$D_g^{(m)} = -\Psi(5 + \frac{1}{6}x^2) \quad (8.41)$$

$$V^{(m)} = \frac{1}{3}\Psi x \quad (8.42)$$

$$\Psi = 3V_0. \quad (8.43)$$

On super-horizon scales, $x \ll 1$ we have

$$D_g^{(r)} \simeq -\frac{20}{3}\Psi \quad \text{and} \quad V^{(r)} \simeq \frac{1}{3}x\Psi\ . \quad (8.44)$$

8.3.2 Isocurvature initial conditions

Here we want to solve the system (7.50) and (7.55) for dark matter and radiation under the condition that the metric perturbations vanish initially, *i.e.*, $\Psi = 0$,

$$\Psi = -\frac{3}{2}\left(\frac{\dot{a}}{a}\right)^2 k^{-2}\left[D_g + 3(1+w)\Psi + 3(1+w)\left(\frac{\dot{a}}{a}\right)k^{-1}V\right] = 0. \quad (8.45)$$

In principle, we have four evolution and one constraint equations. We therefore have four constants to adjust. Condition (8.45), however, requires an entire function to vanish. This may be impossible. Let us nevertheless try:

If $\Psi = 0$ the solutions of the radiation dominated equations are simply

$$D_g^{(r)} = A\cos\left(\frac{x}{\sqrt{3}}\right) + B\sin\left(\frac{x}{\sqrt{3}}\right) \quad (8.46)$$

$$V^{(r)} = \frac{\sqrt{3}}{4}A\sin\left(\frac{x}{\sqrt{3}}\right) - \frac{\sqrt{3}}{4}B\cos\left(\frac{x}{\sqrt{3}}\right). \quad (8.47)$$

For the matter perturbations we find

$$V^{(m)} = -\frac{V_0}{a}, \quad a \propto x^\beta, \quad 1 \le \beta \le 2 \quad (8.48)$$

$$D_g^{(m)} = C^{(m)} - \frac{V_0}{\beta - 1}\frac{x}{a} \quad (8.49)$$

$\Psi = 0$ implies with

$$D_g = \frac{1}{\rho}\left(\rho_r D_g^{(r)} + \rho_m D_g^{(m)}\right) \quad \text{and} \quad (8.50)$$

$$V = \frac{1}{\rho + p}\left((\rho_r + p_r)V^{(r)} + \rho_m V^{(m)}\right) \quad (8.51)$$

that
$$0=\frac{\rho_r}{\rho_m}D_g^{(r)}+D_g^{(m)}+\frac{4\rho_r}{\rho_m}\left(\frac{\dot a}{a}\right)k^{-1}V^{(r)}+3\left(\frac{\dot a}{a}\right)k^{-1}V^{(m)}. \tag{8.52}$$

Since $V^{(m)}\propto 1/a$ it can compensate, for small values of x, the term $\propto\cos(x/\sqrt{3})$ of $V^{(r)}$, which behaves like $1/a$ as well, due to the pre-factor ρ_r/ρ_m. This term can also be compensated in $D_g^{(r)}$ by the term V_0x/a of $D_g^{(m)}$. However, there is no way to compensate $C^{(m)}$ or the term proportional to A. We have to choose therefore $A=C^{(m)}=0$ and

$$a\frac{\rho_r}{\rho_m}\frac{\sqrt{3}}{3}B=V_0,\quad B=\frac{\rho_m}{a\rho_r}\sqrt{3}V_0. \tag{8.53}$$

(The compensation of the smaller terms $D_g^{(r)}$ and $D_g^{(m)}$ is only complete if $\beta\simeq 2$.)

With $c_s=1/\sqrt{3}$ we find

$$D_g^{(r)} \simeq \frac{\rho_m}{a\rho_r c_s}V_0\sin(c_sx)\quad\text{(isocurvature)} \tag{8.54}$$
$$D_g^{(r)} \simeq \Psi\left(8-\frac{44}{3}\cos(c_sx)\right)\quad\text{(adiabatic)} \tag{8.55}$$
$$\tag{8.56}$$

The CMB anisotropies, which we are going to determine in the next chapter, contain a term

$$\frac{\Delta T}{T}(\mathbf{k},\eta_0,\mathbf{n})=\cdots+\frac{1}{4}D_r^{(g)}(\mathbf{k},\eta_{\rm dec})e^{i\mathbf{kn}(\eta_0-\eta_{\rm dec})}\ldots \tag{8.57}$$

On scales where this term dominates the CMB anisotropies, the peaks in D_g translate into peaks in the angular power spectrum of CMB anisotropies.

For isocurvature iuitial conditions, we find a first peak in D_g at

$$x_i^{(1)}=k_i^{(1)}\eta_{\rm dec}=\frac{1}{c_s}\frac{\pi}{2},\quad \lambda_i^{(1)}=\frac{2\pi}{k_i^{(1)}}=4c_s\eta_{\rm dec},\quad \vartheta_i^{(1)}\simeq\frac{4c_s\eta_{\rm dec}}{\chi(\eta_0-\eta_{\rm dec})}, \tag{8.58}$$

Here $\vartheta_i^{(1)}$ is the angle under which the comoving scale $\lambda_i^{(1)}$ at comoving distance $\eta_0-\eta_{\rm dec}$ is seen. In the next chapter, we will expand the temperature fluctuations in terms of spherical harmonics. An fluctuation on the angular scale ϑ then shows up around the harmonic $\ell\sim\pi/(2\vartheta)$. As an indication, we note that for $\Lambda=\kappa=0$, the harmonic of the first isocurvature peak is

$$\ell_i^{(1)}\sim\pi/(2\vartheta_i^{(1)})\sim 110\ ,$$

In the adiabatic case the first "peak" is at $k_a^{(1)}=0$.

Since $D_g^{(r)}$ is negative for small x, the first peaks are "expansion peaks", and due to the gravitational attraction of the baryons (which we have neglected in this simple argument) they are less pronounced than the second ("compression") peaks.

These second peaks are usually called the "first acoustic peak". (It is the first compression peak and we shall adopt the convention to call it the "first peak" mainly for consistency with the literature.) They correspond to wavelengths and angular scales

$$\lambda_i^{(2)} = \frac{4}{3}c_s\eta_{\rm dec}\ ,\quad \vartheta_i^{(2)}\simeq\frac{(4/3)c_s\eta_{\rm dec}}{\chi(\eta_0-\eta_{\rm dec})}\ ,\quad \ell_i^{(2)}\sim 350\quad\text{(isocurvature)} \tag{8.59}$$
$$\lambda_a^{(2)} = 2c_s\eta_{\rm dec}\ ,\quad \vartheta_a^{(2)}\simeq\frac{2c_s\eta_{\rm dec}}{\chi(\eta_0-\eta_{\rm dec})}\ ,\quad \ell_a^{(2)}\sim 220\quad\text{(adiabatic).} \tag{8.60}$$

Here the indicated harmonic is the one obtained in the case $\Lambda=\kappa=0$, for a typical baryon density inferred from nuclosynthesis.

In is interesting to nore that the distance between consecutive peaks is **independent of the initial condition.** It is given by

$$\Delta k_i = k_i^{(2)} - k_i^{(1)} = \pi/(c_s\eta_{\rm dec}) = \Delta k_a\ , \quad \Delta\vartheta = \frac{2c_s\eta_{\rm dec}}{\chi(\eta_0 - \eta_{\rm dec})}\ , \quad \Delta\ell \sim 220\ . \tag{8.61}$$

Again, the numerical value indicated for $\Delta\ell$ corresponds to a universe with $\Lambda = \kappa = 0$. The result is strongly dependent especially on κ. This is the reason why the measurement of the peak position (or better of the inter-peak distance) allows an accurate determination of curvature.

From our analysis we can hence draw the following important conclusions: For scales where this term dominates, the CMB anisotropies show a series of acoustic oscillations with spacing Δk, the position of the first significant peaks is at $k = k_{a/i}^{(2)}$, depending on the initial condition.

The spacing Δk is *independent* of initial conditions. The angle $\Delta\vartheta$ onto which this scale is projected in the sky is determined entirely by the matter content and the geometry of the universe. According to our findings in Chapter I, ϑ will be larger if $\Omega_\kappa < 0$ (positive curvature) and smaller if $\Omega_\kappa > 0$ (see Fig. 6.3).

In our analysis we have neglected the presence of baryons, in order to obtain simple analytical results. Baryons have two effects: They lead to $(\rho+3p)_{\rm rad+bar} > 0$, and therefore to an enhancement of the *compression* peaks (the first, third, etc. acoustic peak). In addition, the baryons slightly *decrease* the sound speed c_s, increasing thereby Δk and decreasing $\Delta\vartheta$.

Another point which we have neglected is the fact that the universe became matter dominated at $\eta_{\rm eq}$, only shortly before decoupling: $\eta_{\rm dec} \simeq 4\eta_{\rm eq}$, for $\Omega_m = 1$. As we have seen, the gravitational potential on sub-horizon scales is decaying in the radiation dominated era. If the radiation dominated era is not very long ago at decoupling, the gravitational potential is still slightly decaying and free streaming photons fall into a deeper gravitational potential than the have to climb out of. This effect, called "early integrated Sachs Wolfe effect" adds to the photon temperature fluctuations at scales which are only slighty larger than the position of the first acoustic peak for adiabatic perturbations. It therefore 'boosts' this peak and, at the same time, moves it to lightly larger scales (smaller angles) Since $\eta_{\rm eq} \propto h^{-2}$, the first acoustic peak is larger if h is smaller.

A small Hubble parameter *increases* therefore the acoustic peaks. A similar effect is observed if a cosmological constant or a negative curvature are present, since $\eta_{\rm eq}$ is retarded in those cases.

8.3.3 Vector perturbations of perfect fluids

If $\Pi^{(V)} = 0$ equation (7.56) implies

$$\Omega \propto a^{3c_s^2-1}. \tag{8.62}$$

For $\dot p/\dot\rho = c_s^2 \le 1/3$, this leads to a non–growing vorticity. The dynamical Einstein equation implies

$$\sigma^{(V)} \propto a^{-2}\ , \tag{8.63}$$

and the constraint (7.51) reads (at early times, so we can neglect curvature)

$$\Omega \sim x^2\sigma^{(V)}. \tag{8.64}$$

If perturbations are created in the very early universe on super–horizon scales (*e.g.* during an inflationary period), vector perturbations of the metric decay and become soon entirely negligible. Even if Ω_i remains constant in a radiation dominated universe, it has to be so small on relevant scales at formation ($x_{in} \ll 1$) that we may safely neglect it.

8.3.4 Tensor perturbations

The situation is different for tensor perturbations. Again we consider the perfect fluid case, $\Pi_{ij}^{(T)} = 0$. There (7.54) implies (if κ is negligible)

$$H_{ij}'' + \frac{2\beta}{x}H_{ij}' + H_{ij} = 0\ , \tag{8.65}$$

with $\beta = 1$ in the radiation dominated era and $\beta = 2$ in the matter dominated era. The less decaying mode solution to Eq. (8.65) is $H_{ij} = e_{ij}x^{1/2-\beta}J_{1/2-\beta}(x)$, where J_ν denotes the Bessel function of order ν and e_{ij} is a transverse traceless polarisation tensor. This leads to

$$H_{ij} = \text{const} \quad \text{for } x \ll 1 \tag{8.66}$$

$$H_{ij} = \frac{1}{a} \quad \text{for } x \gtrsim 1. \tag{8.67}$$

Chapter 9

CMB anisotropies

9.1 Lightlike geodesics

After decoupling, $\eta > \eta_{\rm dec}$, photons follow to a good approximation lightlike geodesics. The temperature shift is then given by the energy shift of a given photon.

The unperturbed photon trajectory follows $(x^\mu) \equiv (\eta, \mathbf{n}(\eta - \eta_0) + \mathbf{x}_0)$, where $\mathbf{x}_0$ is the photon position at time η_0 and $\mathbf{n}$ is the (parallel transported) photon direction. With respect to a geodesic basis $(\mathbf{e})_{i=1}^3$, the components of $\mathbf{n}$ are constant. If $\kappa = 0$ we may choose $\mathbf{e}_i = \partial/\partial x^i$; if $\kappa \neq 0$ these vector fields are no longer parallel transported and therefore do not form a geodesic basis ($\nabla_{\mathbf{e}_i}\mathbf{e}_j = 0$).

Our metric is of the form

$$d\bar{s}^2 = a^2 ds^2 \quad \text{, with} \tag{9.1}$$

$$ds^2 = (\gamma_{\mu\nu} + h_{\mu\nu})\, dx^\mu dx^\nu, \quad \gamma_{00} = -1, \gamma_{i0} = 0, \gamma_{ij} = \gamma_{ji} \tag{9.2}$$

as before.

We make use of the fact that lightlike geodesics are conformally invariant. More precisely ds^2 and $d\bar{s}^2$ have the same lightlike geodesics, only the corresponding affine parameters are different. Let us denote the two affine parameters by $\bar\lambda$ and λ respectively, and the tangent vectors to the geodesic by

$$n = \frac{dx}{d\lambda}, \quad \bar{n} = \frac{dx}{d\bar\lambda} \ , \quad n^2 = \bar{n}^2 = 0 \ , \ n^0 = 1 \ , \ \mathbf{n}^2 = 1. \tag{9.3}$$

We set $n^0 = 1 + \delta n^0$. The geodesic equation for the perturbed metric

$$ds^2 = (\gamma_{\mu\nu} + h_{\mu\nu}) dx^\mu dx^\nu \tag{9.4}$$

yields, to first order,

$$\frac{d}{d\lambda}\delta n^\mu = -\delta\Gamma^\mu_{\alpha\beta} n^\alpha n^\beta. \tag{9.5}$$

For the energy shift, we have to determine δn^0. Since $g^{0\mu} = -1 \cdot \delta_{0\mu}$ + first order, we obtain $\delta\Gamma^0_{\alpha\beta} = -1/2(h_{\alpha 0,\beta} + h_{\beta 0,\alpha} - \dot{h}_{\alpha\beta})$, so that

$$\frac{d}{d\lambda}\delta n^0 = h_{\alpha 0,\beta} n^\beta n^\alpha - \frac{1}{2}\dot{h}_{\alpha\beta} n^\alpha n^\beta. \tag{9.6}$$

Integrating this equation we use $h_{\alpha 0,\beta} n^\beta = \frac{d}{d\lambda}(h_{\alpha 0} n^\alpha)$, so that the change of n^0 between some initial time η_i and some final time η_f is given by

$$\delta n^0|_i^f = \left[h_{00} + h_{0j} n^j\right]_i^f - \frac{1}{2}\int_i^f \dot{h}_{\mu\nu} n^\mu n^\nu d\lambda \ . \tag{9.7}$$

On the other hand, the ratio of the energy of a photon measured by some observer at t_f to the energy emitted at t_i is

$$\frac{E_f}{E_i} = \frac{(\bar{n}\cdot u)_f}{(\bar{n}\cdot u)_i} = \frac{T_f}{T_i}\frac{(n\cdot u)_f}{(n\cdot u)_i}\ , \tag{9.8}$$

where u_f and u_i are the four-velocities of the observer and emitter respectively, and the factor T_f/T_i is the usual (unperturbed) redshift, which relates n and $\bar{n}$. The velocity field of observer and emitter is given by

$$u = (1-A)\partial_\eta + v^i\partial_i\ . \tag{9.9}$$

An observer measuring a temperature T_0 receives photons that were emitted at the time η_{dec} of decoupling of matter and radiation, at the fixed temperature T_{dec}. In first-order perturbation theory, we find the following relation between the unperturbed temperatures T_f, T_i, the measurable temperatures T_0, T_{dec}, and the photon density perturbation:

$$\frac{T_f}{T_i} = \frac{T_0}{T_{dec}}\left(1-\frac{\delta T_f}{T_f}+\frac{\delta T_i}{T_i}\right) = \frac{T_0}{T_{dec}}\left(1-\frac{1}{4}\delta^{(r)}|_i^f\right)\ , \tag{9.10}$$

where $\delta^{(r)}$ is the intrinsic density perturbation in the radiation and we used $\rho^{(r)} \propto T^4$ in the last equality. Inserting the above equation and Eq. (9.7) into Eq. (9.8), and using Eq. (7.19) for the definition of $h_{\mu\nu}$, one finds, after integration by parts [55] the following result for scalar perturbations:

$$\frac{E_f}{E_i} = \frac{T_0}{T_{dec}}\left\{1-\left[\frac{1}{4}D_g^{(r)}+V_j^{(b)}n^j+\Psi-\Phi\right]_i^f+\int_i^f(\dot{\Psi}-\dot{\Phi})d\lambda\right\}\ . \tag{9.11}$$

Here $D_g^{(r)}$ denotes the density perturbation in the radiation fluid, and $V^{(b)}$ is the peculiar velocity of the baryonic matter component (the emitter and observer of radiation). The final time values in the square bracket of Eq. (9.11) give rise only to monopole contributions and to the dipole due to our motion with respect to the CMB, and will be neglected in what follows.

Evaluating Eq. (9.11) at final time η_0 (today) and initial time η_{dec}, we obtain the temperature difference of photons coming from different directions $\mathbf{n}$ and $\mathbf{n}'$

$$\frac{\Delta T}{T} \equiv \frac{\delta T(\mathbf{n})}{T} - \frac{\delta T(\mathbf{n}')}{T}, \tag{9.12}$$

with temperature perturbation

$$\frac{\Delta T(\mathbf{n})}{T} = \left[\frac{1}{4}D_g^{(r)}+V_j^{(b)}n^j+\Psi-\Phi\right](\eta_{dec},\mathbf{x}_{dec})+\int_{\eta_{dec}}^{\eta_0}(\dot{\Psi}-\dot{\Phi})(\eta,\mathbf{x}(\eta))d\eta\ , \tag{9.13}$$

where $\mathbf{x}(\eta) = \mathbf{x}_0-(\eta_0-\eta)\mathbf{n}$ is the unperturbed photon position at time η for an observer at $\mathbf{x}_0$, and $\mathbf{x}_{dec} = \mathbf{x}(\eta_{dec})$. The first term in Eq. (9.13) describes the intrinsic inhomogeneities on the surface of last scattering, due to acoustic oscillations prior to decoupling. Depending on the initial conditions, it can contribute significantly on super-horizon scales. This is especially important in the case of adiabatic initial conditions. As we have seen in Eq. (8.44), in a dust + radiation universe with $\Omega = 1$, adiabatic initial conditions imply $D_g^r(k,\eta) = -20/3\Psi(k,\eta)$ and $V^{(b)} = V^{(r)} \ll D_g^{(r)}$ for $k\eta \ll 1$. With $\Phi = -\Psi$ the the square bracket of Eq. (9.13) gives

$$\left(\frac{\Delta T(\mathbf{n})}{T}\right)^{(OSW)}_{\text{adiabatic}} = \frac{1}{3}\Psi(\eta_{dec},\mathbf{x}_{dec})$$

on super-horizon scales. The contribution to $\frac{\delta T}{T}$ from the last scattering surface on very large scales is called the 'ordinary Sachs Wolfe effect' (OSW). It has been derived for the first time by Sachs and Wolfe [71]. For isocurvature perturbations, the initial condition $D_g(k,\eta) \to 0$ for $\eta \to 0$ is satisfied and the contribution of D_g to the ordinary Sachs Wolfe effect can be neglected.

$$\left(\frac{\Delta T(\mathbf{n})}{T}\right)^{(OSW)}_{\text{isocurvature}} = 2\Psi(\eta_{dec},\mathbf{x}_{dec})$$

The second term in (9.13) describes the relative motions of emitter and observer. This is the Doppler contribution to the CMB anisotropies. It appears on the same angular scales as the acoustic term, and we thus call the sum of the acoustic and Doppler contributions "acoustic peaks".

The last two terms are due to the inhomogeneities in the spacetime geometry; the first contribution determines the change in the photon energy due to the difference of the gravitational potential at the position of emitter and observer. Together with the part contained in $D_g^{(r)}$ they represent the "ordinary" Sachs-Wolfe effect. The integral accounts for red-shift or blue-shift caused by the time dependence of the gravitational field along the path of the photon, and represents the so-called integrated Sachs-Wolfe (ISW) effect. In a $\Omega = 1$, pure dust universe, the Bardeen potentials are constant and there is no integrated Sachs-Wolfe effect; the blue-shift which the photons acquire by falling into a gravitational potential is exactly canceled by the redshift induced by climbing out of it. This is no longer true in a universe with substantial radiation contribution, curvature or a cosmological constant.

The sum of the ordinary Sachs Wolfe term and the integral is the full Sachs-Wolfe contribution (SW).

For **vector** perturbations $\delta^{(r)}$ and A vanish and Eq. (9.8) leads to

$$(E_f/E_i)^{(V)} = (a_i/a_f)[1 - V_j^{(m)} n^j|_i^f + \int_i^f \dot{\sigma}_j n^j d\lambda] \ . \tag{9.14}$$

Again we obtain a Doppler term and a gravitational contribution. For **tensor** perturbations, i.e. gravitational waves, only the gravitational part remains:

$$(E_f/E_i)^{(T)} = (a_i/a_f)[1 - \int_i^f \dot{H}_{lj} n^l n^j d\lambda] \ . \tag{9.15}$$

Equations (9.11), (9.14) and (9.15) are the manifestly gauge invariant results for the Sachs–Wolfe effect for scalar vector and tensor perturbations. Disregarding again the dipole contribution due to our proper motion, Eqs. (9.14,9.15) imply the vector and tensor temperature fluctuations

$$\left(\frac{\Delta T(\mathbf{n})}{T}\right)^{(V)} = V_j^{(m)}(\eta_{dec}, \mathbf{x}_{dec}) n^j + \int_i^f \dot{\sigma}_j(\eta, \mathbf{x}(\eta)) n^j d\lambda \tag{9.16}$$

$$\left(\frac{\Delta T(\mathbf{n})}{T}\right)^{(T)} = -\int_i^f \dot{H}_{lj}(\eta, \mathbf{x}(\eta)) n^l n^j d\lambda] \ . \tag{9.17}$$

Note that for models where initial fluctuations have been led down in the very early universe, vector perturbations are irrelevant as we have aleady pointed out. In this sense Eq. (9.16) is here mainly for completeness. However, in models where perturbations are sourced by some inherently inhomogeneous component (*e.g.* topological defects) vector perturbation can be important.

9.2 Power spectra

One of the basic tools to compare models of large scale structure with observations are power spectra. They are the "harmonic transform" of the two point correlation functions. If the perturbations of the model under consideration are Gaussian (a relatively generic prediction from inflationary models), then the power spectra contain the full statistical information of the model.

One important power spectrum is the dark matter power spectrum,

$$P_D(k) = \left\langle \left| D_g^{(m)}(\mathbf{k}, \eta_0) \right|^2 \right\rangle , \tag{9.18}$$

where $\langle \ \rangle$ indicates a statistical average over "initial conditions" in a given model. $P_D(k)$ is usually compared with the observed power spectrum of the galaxy distribution.

Another power spectrum is given by the velocity perturbations,

$$P_V(k) = \left\langle |\mathbf{V}(\mathbf{k}, \eta_0)|^2 \right\rangle \simeq H_0^2 \Omega^{1.2} P_D(k) k^{-2} \ . \tag{9.19}$$

For $\simeq$ we have used that $|kV|(\eta_0) = \dot{D}_g^{(m)}(\eta_0) \sim H_0\Omega^{0.6}D_g$ on sub-horizon scales (see *e.g.* [60]).

The power spectrum we are most interested in is the CMB anisotropy power spectrum. It is defined as follows: $\Delta T/T$ is a function of position $\mathbf{x}_0$, time η_0 and photon direction $\mathbf{n}$. We develop the $\mathbf{n}$–dependence in terms of spherical harmonics. We will suppress the argument η_0 and often also $\mathbf{x}_0$ in the following calculations. All results are for today (η_0) and here ($\mathbf{x}_0$). By statistical homogeneity expectation values are supposed to be independent of position.

$$\frac{\Delta T}{T}(\mathbf{x}_0, \mathbf{n}, \eta_0) = \sum_{\ell,m} a_{\ell m}(\mathbf{x}_0)Y_{\ell m}(\mathbf{n}), \quad \langle a_{\ell m} \cdot a^*_{\ell' m'}\rangle = \delta_{\ell\ell'}\delta_{mm'}C_\ell \tag{9.20}$$

The C_ℓ's are the CMB power spectrum. We assume that the perturbations are generated by a ***homogeneous*** and ***isotropic*** process, so that C_ℓ depends neither on $\mathbf{x}_0$ nor on m, and that $\langle a_{\ell m} \cdot a^*_{\ell' m'}\rangle$ vanishes for $\ell \neq \ell'$ or $m \neq m'$.

Let us, at this point insert a comment on the problem of **cosmic variance**: Even if our 'ergodic hypothesis' is correct and we may interchange ensemble and spacial averages, we cannot obtain very precice averages for measurements of large scale characteristics, due to the fact that we can observe only the universe around a given position. For example, let us assume that temperature fluctuations are Gaussian, as they are in most inflationary models. The functions $a_{\ell m}$ are then also Gaussian distributed, and we expect a variance of

$$|\frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell}|a_{\ell m}|^2 - C_\ell| = |C_\ell^{obs} - C_\ell| = \frac{C_\ell}{2\ell+1} ,$$

on the average of the $2\ell+1$ values $a_{\ell m}$ which can in principle be measured from one point. For simplicity, we neglect here the additional reduction due to the fact that our own milky ways blocks a portion of sky (about 20%). Wick's theorem now gives

$$\langle C_\ell^2\rangle - \langle C_\ell\rangle^2 = \langle|a_{\ell m}|^4\rangle - \langle|a_{\ell m}|^2\rangle^2 = 2\langle|a_{\ell m}|^2\rangle^2$$

For a given multipole ℓ we then expect a variance of

$$\frac{\sqrt{(C_\ell^{obs})^2 - C_\ell^2}}{C_\ell} = \sqrt{\frac{2}{2\ell+1}} . \tag{9.21}$$

The two point correlation function is related to the C_ℓ's by

$$\left\langle\frac{\Delta T}{T}(\mathbf{n})\frac{\Delta T}{T}(\mathbf{n}')\right\rangle_{\mathbf{n}\cdot\mathbf{n}'=\mu} = \sum_{\ell,\ell',m,m'}\langle a_{\ell m} \cdot a^*_{\ell' m'}\rangle Y_{\ell m}(\mathbf{n})Y^*_{\ell m}(\mathbf{n}') =$$

$$\sum_\ell C_\ell \underbrace{\sum_{m=-\ell}^{\ell} Y_{\ell m}(\mathbf{n})Y^*_{\ell m}(\mathbf{n}')}_{\frac{2\ell+1}{4\pi}P_\ell(\mathbf{n}\cdot\mathbf{n}')} = \frac{1}{4\pi}\sum_\ell(2\ell+1)C_\ell P_\ell(\mu), \tag{9.22}$$

where we have used the addition theorem of spherical harmonics for the last equality.

Clearly the a_{lm}'s from scalar, vector and tensor perturbations are uncorrelated,

$$\left\langle a^{(S)}_{\ell m}a^{(V)}_{\ell' m'}\right\rangle = \left\langle a^{(S)}_{\ell m}a^{(T)}_{\ell' m'}\right\rangle = \left\langle a^{(V)}_{\ell m}a^{(T)}_{\ell' m'}\right\rangle = 0. \tag{9.23}$$

Since vector perturbations decay, their contributions, the $C_\ell^{(V)}$, are negligible in models where initial perturbations have been laid down very early, *e.g.*, after an inflationary period. Tensor perturbations are constant on super-horizon scales and perform damped oscillations once they enter the horizon.

Let us first discuss in somewhat more detail scalar perturbations. We suppose the initial perturbations to be given by a spectrum,

$$\left\langle|\Psi|^2\right\rangle k^3 = Ak^{n-1}\eta_0^{n-1}. \tag{9.24}$$

(We multiply by the constant η_0^{n-1} in order to keep A dimensionless for all values of n.)

On *super-horizon scales* we then have, for *adiabatic* perturbations:

$$\frac{1}{4}D_g^{(r)} = -\frac{5}{3}\Psi + \mathcal{O}(x^2), \quad V^{(b)} = V^{(r)} = \mathcal{O}(x) \tag{9.25}$$

The dominant contribution on super-horizon scales (neglecting the integrated Sachs Wolfe effect $\int \dot{\Phi} - \dot{\Psi}$) is then

$$\frac{\Delta T}{T}(\mathbf{x}_0, \mathbf{n}, \eta_0) = \frac{1}{3}\Psi(x_{\rm dec}, \eta_{\rm dec}). \tag{9.26}$$

The Fourier transform of (9.26) gives (setting $\mu \equiv \widehat{\mathbf{k}}\mathbf{n}$)

$$\frac{\Delta T}{T}(\mathbf{k}, \mathbf{n}, \eta_0) = \frac{1}{3}\Psi(k, \eta_{\rm dec}) \cdot e^{i\mathbf{kn}(\eta_0 - \eta_{\rm dec})} . \tag{9.27}$$

Using

$$e^{i\mathbf{kn}(\eta_0 - \eta_{\rm dec})} = \sum_{\ell=0}^{\infty}(2\ell+1)i^\ell j_\ell(k(\eta_0 - \eta_{\rm dec}))P_\ell(\mu) ,$$

we obtain

$$\begin{aligned}
&\left\langle \frac{\Delta T}{T}(\mathbf{x}_0, \mathbf{n}, \eta_0)\frac{\Delta T}{T}(\mathbf{x}_0, \mathbf{n}', \eta_0)\right\rangle \\
&= \frac{1}{V}\int d^3x_0 \left\langle \frac{\Delta T}{T}(\mathbf{x}_0, \mathbf{n}, \eta_0)\frac{\Delta T}{T}(\mathbf{x}_0, \mathbf{n}', \eta_0)\right\rangle \\
&= \frac{1}{(2\pi)^3}\int d^3k \left\langle \frac{\Delta T}{T}(\mathbf{k}, \mathbf{n}, \eta_0)\left(\frac{\Delta T}{T}\right)^*(\mathbf{k}, \mathbf{n}', \eta_0)\right\rangle \\
&= \frac{1}{(2\pi)^3 9}\int d^3k \left\langle |\Psi|^2\right\rangle \sum_{\ell,\ell'=0}^{\infty}(2\ell+1)(2\ell'+1)j_\ell(k(\eta_0 - \eta_{\rm dec}))j'_\ell(k(\eta_0 - \eta_{\rm dec}))i^{\ell-\ell'} \\
&\cdot P_\ell(\hat{\mathbf{k}}\mathbf{n}) \cdot P'_\ell(\hat{\mathbf{k}}\mathbf{n}') .
\end{aligned} \tag{9.28}$$

Inserting $P_\ell(\hat{\mathbf{k}}\mathbf{n}) = \frac{4\pi}{2\ell+1}\sum_m Y^*_{\ell m}(\hat{\mathbf{k}})Y_{\ell m}(\mathbf{n})$ and $P'_\ell(\hat{\mathbf{k}}\mathbf{n}') = \frac{4\pi}{2\ell'+1}\sum_{m'} Y^*_{\ell' m'}(\hat{\mathbf{k}})Y_{\ell' m'}(\mathbf{n}')$, integration over the directions $d\Omega_{\hat{k}}$ gives $\delta_{\ell\ell'}\delta_{mm'}\sum_m Y^*_{\ell m}(\mathbf{n})Y_{\ell m}(\mathbf{n}')$. Using as well $\sum_m Y^*_{\ell m}(\mathbf{n})Y_{\ell m}(\mathbf{n}') = \frac{2\ell+1}{4\pi}P_\ell(\mu)$, where now $\mu = \mathbf{n}\cdot\mathbf{n}'$, we find

$$\begin{aligned}
&\left\langle \frac{\Delta T}{T}(\mathbf{x}_0, \mathbf{n}, \eta_0)\frac{\Delta T}{T}(\mathbf{x}_0, \mathbf{n}', \eta_0)\right\rangle_{\mathbf{nn}'=\mu} = \\
&\sum_\ell \frac{2\ell+1}{4\pi}P_\ell(\mu)\frac{2}{\pi}\int\frac{dk}{k}\left\langle\frac{1}{9}|\Psi|^2\right\rangle k^3 j_\ell^2(k(\eta_0 - \eta_{\rm dec})).
\end{aligned} \tag{9.29}$$

Comparing this equation with Eq. (9.22) we obtain for *adiabatic perturbations* on scales $2 \le \ell \ll \chi(\eta_0 - \eta_{\rm dec})/\eta_{\rm dec} \sim 100$

$$C_\ell^{(SW)} \simeq C_\ell^{(OSW)} \simeq \frac{2}{\pi}\int_0^\infty \frac{dk}{k}\left\langle\left|\frac{1}{3}\Psi\right|^2\right\rangle k^3 j_\ell^2\left(k\left(\eta_0 - \eta_{\rm dec}\right)\right). \tag{9.30}$$

If Ψ is a pure power law and we set $k(\eta_0 - \eta_{\rm dec}) \sim k\eta_0$, the integral (9.30) can be performed analytically. For the ansatz (9.24) one obtains for $-3 < n < 3$

$$C_\ell^{(SW)} = \frac{A}{9}\frac{\Gamma(3-n)\Gamma(\ell - \frac{1}{2} + \frac{n}{2})}{2^{3-n}\Gamma^2(2 - \frac{n}{2})\Gamma(\ell + \frac{5}{2} - \frac{n}{2})}. \tag{9.31}$$

Of special interest is the *scale invariant* spectrum, $n = 1$. This spectrum with a time and scale independent gravitational potential has first been investigated by Harrison and by Zel'dovich [73]. It is therefore called the Harrison–Zel'dovich spectrum and leads to

$$\ell(\ell+1)C_\ell^{(SW)} = \text{const.} \simeq \left\langle\left(\frac{\Delta T}{T}(\vartheta_\ell)\right)^2\right\rangle , \quad \vartheta_\ell \equiv \pi/\ell . \tag{9.32}$$

This is precicely (within the accuracy of the experiment) the behaviour observed by the DMR experiment aboard COBE [57].

Inflationary models predict very generically a HZ spectrum (up to logarithmic corrections). The DMR discovery has therefore been regarded as a great success, if not a proof, of inflation. There are however other models like topological defects [75, 76, 77] or certain string cosmology models [78] which also predict scale–invariant, *i.e.*, Harrison Zel'dovich spectra of fluctuations. These models do however not belong to the class investigated here, since in these models perturbations are induced by seeds which evolve non–linearly in time.

For isocurvature perturbations, the main contribution on large scales comes from the integrated Sachs Wolfe effect and (9.30) is replaced by

$$C_\ell^{(ISW)} \simeq \frac{2}{\pi}\int\frac{dk}{k}k^3\left\langle\left|\int_{\eta_{\rm dec}}^{\eta_0}2\dot\Psi(k,\eta)j_\ell^2(k(\eta_0-\eta))d\eta\right|^2\right\rangle. \tag{9.33}$$

Inside the horizon Ψ is roughly constant (matter dominated). Using the ansatz (9.24) for Ψ inside the horizon and setting the integral in (9.33) $\sim 2\Psi(k,\eta=1/k)j_\ell^2(k\eta_0)$, we obtain again (9.31), but with $A/9$ replaced by $4A$. The Sachs Wolfe temperature anisotropies coming from isocurvature perturbations are therefore about a factor of 6 times larger than those coming from adiabatic perturbations.

On smaller scales, $\ell \gtrsim 100$ the contribution to $\Delta T/T$ is usually dominated by acoustic oscillations, the first two terms in Eq. (9.13). Instead of (9.33) we then obtain

$$C_\ell^{(AC)} \simeq \frac{2}{\pi}\int_0^\infty\frac{dk}{k}k^3\left\langle\left|\frac{1}{4}D_g^{(r)}(k,\eta_{\rm dec})j_\ell(k\eta_0)+V^{(r)}(k,\eta_{\rm dec})j_\ell'(k\eta_0)\right|^2\right\rangle. \tag{9.34}$$

On sub-horizon scales $D_g^{(r)}$ and $V^{(r)}$ are oscillating like sine or cosine waves depending on the initial conditions. Correspondingly the $C_\ell^{(AC)}$ will show peaks and minima. On very small scales they are damped by the photon diffusion which takes place during the recombination process (see next section).

For gravitational waves (tensor fluctuations), a formula analogous to (9.31) can be derived (see appendix),

$$C_\ell^{(T)} = \frac{2}{\pi}\int dk k^2\left\langle\left|\int_{\eta_{\rm dec}}^{\eta_0}d\eta\dot H(\eta,k)\frac{j_\ell(k(\eta_0-\eta))}{(k(\eta_0-\eta))^2}\right|^2\right\rangle\frac{(\ell+2)!}{(\ell-2)!}. \tag{9.35}$$

To a very crude approximation we may assume $\dot H=0$ on super-horizon scales and $\int d\eta\dot H j_\ell(k(\eta_0-\eta))\sim H(\eta=1/kj_\ell(k\eta_0)$. For a pure power law,

$$k^3\left\langle|H(k,\eta=1/k)|^2\right\rangle = A_T k^{n_T}\eta_0^{-n_T}, \tag{9.36}$$

this gives

$$\begin{aligned} C_\ell^{(T)} &\simeq \frac{2}{\pi}\frac{(\ell+2)!}{(\ell-2)!}A_T\int\frac{dx}{x}x^{n_T}\frac{j_\ell^2(x)}{x^4} \\ &= \frac{(\ell+2)!}{(\ell-2)!}A_T\frac{\Gamma(6-n_T)\Gamma(\ell-2+\frac{n_T}{2})}{2^{6-n_T}\Gamma^2(\frac{7}{2}-n_T)\Gamma(\ell+4-\frac{n_T}{2})}. \end{aligned} \tag{9.37}$$

For a scale invariant spectrum ($n_T=0$) this results in

$$\ell(\ell+1)C_\ell^{(T)} \simeq \frac{\ell(\ell+1)}{(\ell+3)(\ell-2)}. \tag{9.38}$$

The singularity at $\ell=2$ in this crude approximation is not real, but there is some enhancement of $\ell(\ell+1)C_\ell^{(T)}$ at $\ell\sim 2$.

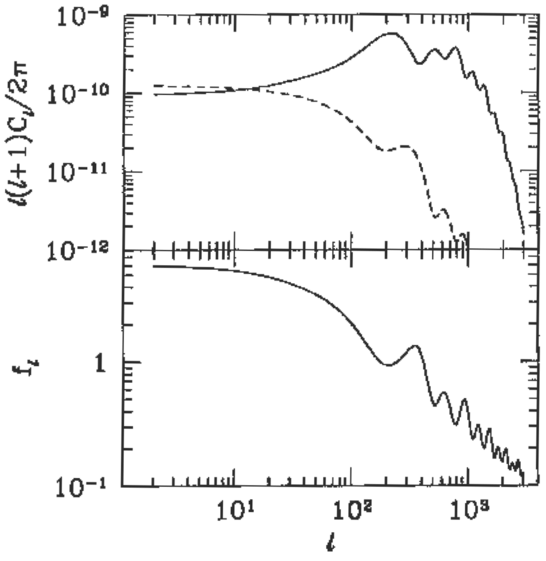

Figure 9.1: A sample adiabatic(solid line) and isocurvature (dashed line) CMB anisotropy spectrum, $\ell(\ell+1)C_\ell$, are shown on the top panel. The quantity shown in the bottom panel is proportional to their ratio (from Kanazawa et al. [79]).

Since tensor perturbations decay on sub-horizon scales, $\ell \gtrsim 60$, they are not very sensitive to cosmological parameters.

Again, inflationary models (and topological defects) predict a scale invariant spectrum of tensor fluctuations ($n_T \sim 0$).

On very small angular scales, $\ell \gtrsim 600$, fluctuations are damped by collisional damping (Silk damping). This effect has to be discussed with the Boltzmann equation for photons derived in the next section.

9.3 The Boltzmann equation

9.3.1 Elements of the derivation

When particles are not very tightly coupled, the fluid approximation breaks down and they have to be described by a Boltzmann equation,

$$p^\mu \partial_\mu f - \Gamma^i_{\alpha\beta} p^\alpha p^\beta \frac{\partial f}{\partial p^i} = C[f] \ . \tag{9.39}$$

$C[f]$ is a collision integral which describes the interactions with other matter components. The left hand side of (9.39) just requires the particles to move along geodesics in the absence of collisions.

Let us first consider the situation where collisions are negligible, $C[f] = 0$. The unperturbed Boltzmann equation implies that f be a function of $v = ap$ only. Setting $f = \bar{f}(v) + F(\eta, \mathbf{x}, v, \mathbf{n})$, where $\mathbf{n}$ denotes the momentum directions, leads then to the perturbation equation

$$\partial_\eta F - n^i \partial_i F - \Gamma^{(S)\,i}_{jk} n^j n^k \frac{\partial F}{\partial n^i} = v \frac{d\bar{f}}{dv} \left[n^i A_{,i} - n^i n^j \left(B_{i|j} - \dot{H}_{ij} \right) + H_L \right] . \tag{9.40}$$

Here $\Gamma^{(S)\,i}_{jk}$ are the Christoffel symbols of the space of constant curvature κ.

To derive (9.40) we have used $p^2 = 0$. The Liouville equation for particles with non–vanishing mass can be found in Ref. [55].

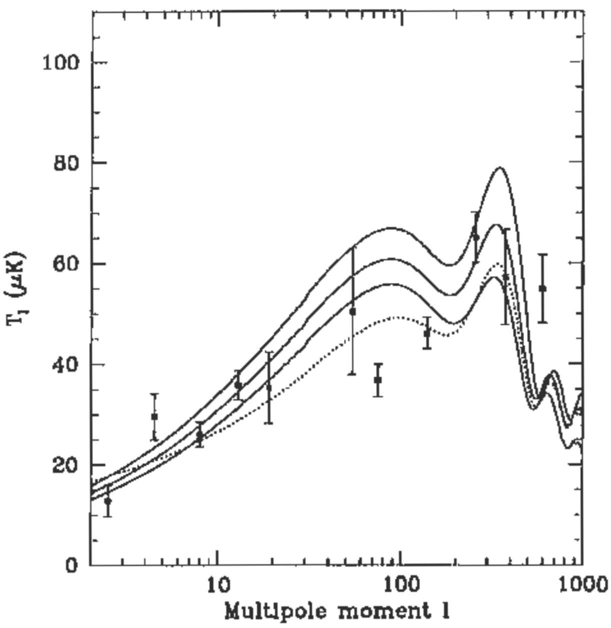

Figure 9.2: Some sample isocurvature CMB anisotropy spectra are shown. The variable $T_\ell = T_0\sqrt{\ell(\ell+1)C_\ell/(2\pi)}$ is plotted for the Peebles model (from [80]).

The ansatz

$$f(x,\mathbf{p}) = \bar{f}\left(\frac{g^{(3)}(\mathbf{p},\mathbf{p})^{\frac{1}{2}}}{T(x,\mathbf{n})}\right) = \bar{f}\left(\frac{v}{T(x,\mathbf{n})}\right) \tag{9.41}$$

with $T(x,\mathbf{n}) = \overline{T}(\eta) + \Delta T(x,\mathbf{n})$ leads to

$$f = \bar{f} - v\frac{d\bar{f}}{dv}\frac{\Delta T}{\overline{T}}. \tag{9.42}$$

Integrating (9.41) over photon energies, we can also write

$$\frac{\Delta T}{T} = \frac{1}{4}\imath, \tag{9.43}$$

where $\imath$ is the brightness perturbation,

$$\imath = \frac{4\pi}{\bar{\rho}a^4}\int_0^\infty Fv^3 dv. \tag{9.44}$$

Comparing Eq. (9.42) with (9.40), we find

$$\partial_\eta\left(\frac{\Delta T}{T}\right) + n^i\partial_i\left(\frac{\Delta T}{T}\right) - \Gamma^{(S)\,i}_{jk}n^j n^k\frac{\partial\left(\frac{\Delta T}{T}\right)}{\partial n^i} = -\left[n^i A_{,i} - \left(B_{i|j} - \dot{H}_{ij}\right)n^i n^j + H_L\right]. \tag{9.45}$$

The fact that gravitational perturbations of Liouville's equation can be cast purely in *temperature perturbations* of the original distribution is not astonishing. This is just an expression of gravity being "achromatic", *i.e.* independent of the photon energy.

We now decompose (9.45) into scalar, vector and tensor components. Even though $\Delta T/T$ is just a function, it can be represented in the form

$$\frac{\Delta T}{T}(\mathbf{x},\mathbf{n}) = \sum_{\ell=0}^\infty \alpha_{i_1,\ldots,i_\ell}(\mathbf{x})n^{i_1}\cdots n^{i_\ell}, \tag{9.46}$$

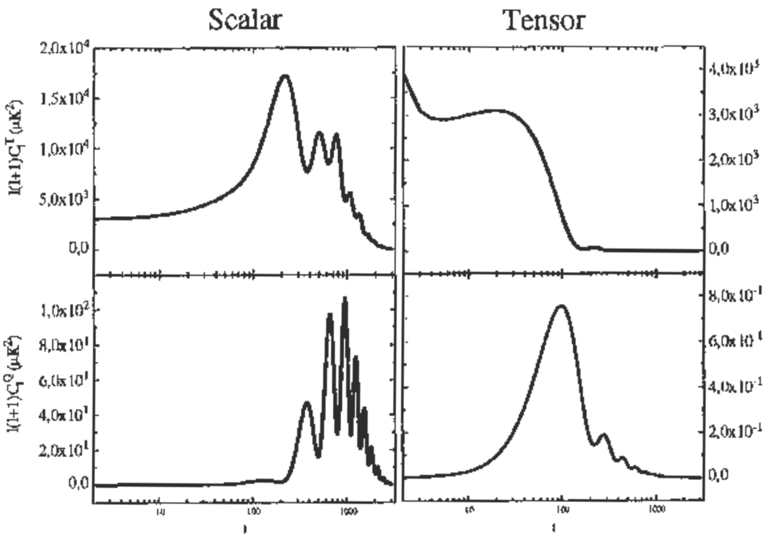

Figure 9.3: Adiabatic scalar and tensor CMB anisotropy spectra are shown (top panels). The bottom panels show the corresponding posarization spectra (see Section 9.4). (from [81]).

where the $\alpha_{i_1,\dots,i_\ell}$ are symmetric traceless tensor fields that contain scalar, vector, 2–tensor and in principle also higher tensor components. Since spin components larger than 2 are not sourced by the right hand side of equation (9.45) and since they are suppressed at early times, when collisions are important, we neglect them here.

For the *scalar* contribution to $\Delta T/T$ we obtain from (9.45)

$$\partial_\eta \left(\frac{\Delta T}{T}\right)^{(S)} + k\mu \left(\frac{\Delta T}{T}\right)^{(S)} - \Gamma^{(S)\,i}_{jk} n^j n^k \frac{\partial \left(\frac{\Delta T}{T}\right)^{(S)}}{\partial n^i} =$$
$$- \left[k\mu A + m u^2 k^2 \left(B - \dot{H}_T\right) + H_L + \frac{1}{3} k^2 \dot{H}_T\right], \qquad (9.47)$$

where we have introduced the direction cosine μ defined by $n^i Y_{,i} = k\mu Y$. Note that in flat space, $\kappa = 0$, we have just $\mu = i\hat{k} \cdot n$.

This equation is not manifestly gauge–invariant. However, setting

$$\mathcal{M} = \left(\frac{\Delta T}{T}\right)^{(S)} + H_L + \frac{1}{3} k^2 H_T + k\mu \left(\dot{H} - B\right), \qquad (9.48)$$

it reduces to

$$\partial_\eta \mathcal{M} + k\mu \mathcal{M} - \Gamma^{(S)\,i}_{jk} n^j n^k \frac{\partial \mathcal{M}}{\partial n^i} = k\mu \left(\Phi - \Psi\right), \qquad (9.49)$$

where Φ and Ψ are the Bardeen potentials. If n^j are components with respect to a geodesic basis (or $\kappa = 0$), the third term on the left hand side of Eq. (9.49) vanishes. For simplicity we now concentrate on the case $\kappa = 0$. We can then integrate the equation and obtain

$$\begin{aligned} \mathcal{M}(\eta_0, \mathbf{n}, \mathbf{k}) &= \exp[i\mathbf{k} \cdot \mathbf{n}(\eta_{in} - \eta_0)] \mathcal{M}(\eta_{in}, \mathbf{n}, \mathbf{k}) \\ &\quad + \int_{\eta_{in}}^{\eta_0} i \exp[i\mathbf{k} \cdot \mathbf{n}(\eta - \eta_0)] \mathbf{n} \cdot \mathbf{k} \left(\Phi - \Psi\right) d\eta \,. \end{aligned} \qquad (9.50)$$

Integration by parts and neglecting the monopole term $(\Phi - \Psi)(\eta_0)$, leads to

$$\mathcal{M}(\eta_0, \mathbf{n}, \mathbf{k}) =$$

$$\exp[i\mathbf{k}\cdot\mathbf{n}(\eta_{in}-\eta_0)]\left[\mathcal{M}(\eta_{in},\mathbf{n},\mathbf{k})+(\Phi-\Psi)(\eta_{in},\mathbf{k})\right] \\ -\int_{\eta_{in}}^{\eta_0}\exp[i\mathbf{k}\cdot\mathbf{n}(\eta-\eta_0)]\left(\dot\Phi-\dot\Psi\right)d\eta \;. \tag{9.51}$$

Comparing this equation with (9.13), we see again that $\mathcal{M}=\left(\frac{\Delta T}{T}\right)^{(S)}$ (up to gauge dependent monopole and dipole contributions) if the initial condition is

$$\mathcal{M}(\eta_{in},\mathbf{n},\mathbf{k})=\frac{1}{4}D_g^{(r)}(\eta_{in},\mathbf{k})+\mathbf{n}\cdot\mathbf{k}V^{(b)}(\eta_{in},\mathbf{k})\;,$$

which is equivalent to require that $\mathcal{M}(\eta_{in})$ has no higher than first moments. As we shall see below, this assumption is quite reasonable since collisions suppress the build up of higher moments before recombination.

Since the right hand side of (9.49) is gauge invariant, the left hand side must be so as well and we conclude that $\mathcal{M}$ is a gauge–invariant variable (a direct proof of this, analysing the gauge transformation properties of the distribution function, can be found in Ref. [55]).

$\mathcal{M}$ coincides with the scalar temperature fluctuations up a to a gauge dependent monopole and dipole contribution.

The vector and tensor parts of $\Delta T/T$ are gauge–invariant by themselves and we denote them by $\mathcal{M}^{(V)}$ and $\mathcal{M}^{(T)}$ for consistency. In the absence of collisions, they satisfy the equations

$$\dot{\mathcal{M}}^{(V)}+i\mu k\mathcal{M}^{(V)}-\Gamma^{(S)\,i}_{jk}n^jn^k\frac{\partial\mathcal{M}^{(V)}}{\partial n^i} = -in^\ell n^m k_\ell\sigma_m^{(V)} \tag{9.52}$$

$$\dot{\mathcal{M}}^{(T)}+i\mu k\mathcal{M}^{(T)}-\Gamma^{(S)\,i}_{jk}n^jn^k\frac{\partial\mathcal{M}^{(T)}}{\partial n^i} = -in^\ell n^m \dot H_{m\ell}. \tag{9.53}$$

The components of the energy momentum tensor are obtained by integrating the second moments of the distribution function over the mass shell,

$$T^{\mu\nu}=\int_{P_m(x)}p^\mu p^\nu f(p,x)\frac{p^2dpd\Omega_{\hat{\mathbf{p}}}}{p^0}\;, \tag{9.54}$$

where $\Omega_{\hat{\mathbf{p}}}$ denotes the angular integration over momentum directions. One finds for $\kappa=0$ and setting $\mu=\mathbf{n}\cdot\hat{\mathbf{k}}$:

$$D_g^{(r)} = \frac{1}{\pi}\int\mathcal{M}d\Omega \tag{9.55}$$

$$V^{(r)} = \frac{3i}{4\pi}\int\mu\mathcal{M}d\Omega \tag{9.56}$$

$$\Pi^{(r)} = \frac{9}{2\pi}\int\left(\mu^2-\frac{1}{3}\right)\mathcal{M}d\Omega \tag{9.57}$$

$$\Gamma^{(r)} = 0 \tag{9.58}$$

$$V_i^{(V)} = \frac{1}{4\pi}\int n_i\mathcal{M}^{(V)}d\Omega \tag{9.59}$$

$$\Pi_j^{(V)} = \frac{6}{\pi}\int\mu n_j\mathcal{M}^{(V)}d\Omega \tag{9.60}$$

$$\Pi_{ij}^{(T)} = \frac{3}{\pi}\int n_in_j\mathcal{M}^{(T)}d\Omega. \tag{9.61}$$

Let us now turn to the collision term. At recombination (when the fluid description of radiation breaks down) the temperature is ~ 0.4 eV. The electrons and nuclei are non–relativistic and the dominant collision process is non–relativistic Thomson scattering. Since collisions are important only before and during recombination, where curvature effects are entirely negligible, we set $\kappa=0$ in the reminder of this section.

The collision term is given by

$$C[f] = \frac{df_+}{d\eta} - \frac{df_-}{d\eta}, \tag{9.62}$$

where f_+ and f_- denote the distribution of photons scattered into respectively out of the beam due to Compton scattering.

In the matter (baryon/electron) rest frame, which we indicate by a prime, we know

$$\frac{df'_+}{dt'}(p, \mathbf{n}) = \frac{\sigma_T n_e}{4\pi} \int f'(p', \mathbf{n}')\omega(\boldsymbol{n}, \boldsymbol{n}')d\Omega' ,$$

where n_e denotes the number density of free electrons, σ_T is the Thomson cross section, and ω is the normalized angular dependence of the Thomson cross section:

$$\omega(\boldsymbol{n}, \boldsymbol{n}') = 3/4[1 + (\boldsymbol{n} \cdot \boldsymbol{n}')^2] = 1 + \frac{3}{4} n_{ij} n'_{ij} \quad \text{with} \quad n_{ij} \equiv n_i n_j - \frac{1}{3}\delta_{ij} .$$

In the baryon rest frame which moves with four velocity u, the photon energy is given by

$$p' = p_\mu u^\mu .$$

We denote by p the photon energy with respect to a tetrad adapted to the slicing of spacetime into $\{\eta = \text{constant}\}$ hyper-surfaces:

$$p = p_\mu n^\mu , \qquad \text{with} \quad n = a^{-1}[(1 - A)\partial_\eta + B^i \partial_i] .$$

The unit vector n is the normal the the hyper-surfaces of constant time. The lapse function and the shift vector of the slicing are given by $\alpha = a(1 + A)$ and $\beta = -B^i \partial_i$. In first order,

$$p_0 = ap(1 + A) - apn_i B^i \quad ,$$

and in zeroth order, clearly,

$$p_i = apn_i .$$

Furthermore, the baryon four velocity has the form $u^0 = a^{-1}(1 - A) \quad , \quad u^i = u^0 v^i$ up to first order. This yields

$$p' = p_\mu u^\mu = p(1 + n_i(v^i - B^i)) .$$

Using this identity and performing the integration ovcr photon energies, we find

$$\rho_r \frac{d\iota_+(n)}{dt'} = \rho_r \sigma_T n_e [1 + 4n_i(v^i - B^i) + \frac{1}{4\pi} \int \iota(n')\omega(n, n')d\Omega'] .$$

The distribution of photons scattered out of the beam, has the well known form (see e.g. Lifshitz and Pitajewski [1983])

$$\frac{df_-}{dt'} = \sigma_T n_e f'(p', \mathbf{n}) ,$$

so that we finally obtain

$$C' = \frac{4\pi}{\rho_r a^4} \int dp (\frac{df_+}{dt'} - \frac{df_-}{dt'}) p^3 = \sigma_T n_e [\delta_r - \iota + 4n_i(v^i - B^i) + \frac{3}{16\pi} n_{ij} \int \iota(n') n'_{ij} d\Omega'] ,$$

where $\delta_r = (1/4\pi) \int \iota(n) d\Omega$ is the photon energy density perturbation.
Using the definitions of the gauge–invariant variables $\mathcal{M}$ and $V^{(b)}$ for the photon brightness perturbation and the baryon velocity potential, we can write C' in gauge–invariant form.

$$C' = 4\sigma_T n_e [\frac{1}{4} D_g^{(r)} - \mathcal{M} + n^i V_i^{(b)} + \frac{1}{2} n_{ij} M^{ij}] , \tag{9.63}$$

with $D_g^{(r)} = (1/\pi)\int \mathcal{M} d\Omega$ and

$$M^{ij} \equiv \frac{3}{8\pi}\int \mathcal{M}(n')n'_{ij}d\Omega' \ .$$

Since the term in square brackets of (9.63) is already first order we can set $dt' = dt$ which yields $C = \frac{dt'}{d\eta}C' = \frac{dt}{d\eta}C' = aC''$. The Boltzmann equation for scalar perturbations expressed in terms of the gauge invariant variable $\mathcal{M}$ then becomes

$$\dot{\mathcal{M}} + n^i\partial_i\mathcal{M} = n^i\partial_i(\Phi - \Psi) + a\sigma_T n_e[\frac{1}{4}D_g^{(r)} - \mathcal{M} - n^i\partial_i V^{(b)} + \frac{1}{2}n_{ij}M^{ij}] \ . \tag{9.64}$$

For vector and tensor perturbations we get

$$\dot{\mathcal{M}}^{(V)} + i\mu k\mathcal{M}^{(V)} = -n^i n^j\sigma_{i|j} + a\sigma_T n_e\left[n^i V_i^{(Vb)} + \frac{1}{2}n^{ij}M_{ij}^{(V)} - \mathcal{M}^{(V)}\right] \tag{9.65}$$

$$\dot{\mathcal{M}}^{(T)} + i\mu k\mathcal{M}^{(T)} = -n^i n^j\dot{H}_{ij} + a\sigma_T n_e\left[n^{ij}M_{ij}^{(T)} - \mathcal{M}^{(T)}\right] \ . \tag{9.66}$$

9.3.2 The tight coupling limit

Before recombination, when $n_e \simeq \rho_B/m_p$,

$$\eta_T \equiv \frac{1}{a\sigma_T n_e} \sim \frac{8}{\Omega_B h}(1+z)^{-\frac{3}{2}}\eta \ll \eta, \quad z \gtrsim z_{\rm dec}. \tag{9.67}$$

To lowest order in η_T, collisions force the photon distribution to be of the form

$$\mathcal{M} = \frac{1}{4}D_g + n^i V_i^{(b)} + \frac{1}{2}n^{ij}M_{ij}, \tag{9.68}$$

the building up of higher moments is strongly suppressed by collisions.

During recombination, the number density of free electrons, n_e, decreases rapidly and the collision term becomes less and less important. Higher moments in the photon distribution develop by free streaming.

The collision term $C[\mathcal{M}]$ of equation (9.64) also appears in the equation of motion of the baryons as a drag. The Thomson drag force is given by

$$F_j = \frac{\rho_r}{4\pi}\int C[\mathcal{M}]n_j d\Omega = \frac{-4a\sigma_T n_e\rho_r}{3}(M_j + V_i^{(b)}) \ , \tag{9.69}$$

$$\text{with} \quad M_j = \frac{3}{4\pi}\int n_j\mathcal{M}d\Omega \ .$$

This yields the following scalar baryon equation of motion in an ionized plasma

$$\partial_j\dot{V}^{(b)} + (\dot{a}/a)\partial_i V^{(b)} = \partial_i\Psi - \frac{4a\sigma_T n_e\rho_r}{3\rho_b}(M_j + \partial_i V^{(b)}) \ , \tag{9.70}$$

where we have added the drag force to the second eq. of (7.55) with $w = c_s^2 = 0$.

We now want to discuss equations (9.64,9.70) in the limit of very many collisions. The comoving photon mean free path is given by $\eta_T = l_T = (a\sigma_T n_e)^{-1}$. In lowest order η_T/η and l_T/λ, [1] $\mathcal{M}$ is given by (9.68), and eq. (9.70) implies

$$M_j + \partial_i V^{(b)} = 0 \ .$$

Inserting the solution (9.68) in the Boltzmann equation (9.64) and integrating over directions this implies

$$\triangle V^{(b)} = \partial_i M_i = \triangle V^{(r)} = \frac{3}{4}\dot{D}_g^{(r)} \ , \tag{9.71}$$

[1] Here λ is a typical size of a perturbation. For a given Fourier mode k, it is $\lambda \sim \pi/k$.

Implying especially $V^{(b)} = V^{(r)} \equiv V$. Eq. (9.71) is equivalent to the energy conservation equation (7.55) for radiation. Using also (7.55) for baryons, $w = 0$, we obtain

$$\dot{D}_g^{(r)} = \frac{4}{3}\triangle V^{(b)} = \frac{4}{3}\dot{D}_g^{(b)}.$$

This shows that entropy per baryon is conserved, $\Gamma = 0$. Before recombination, when the collisions are sufficiently frequent, baryons and photons are adiabatically coupled. Inserting (9.68) in (9.64) we find up to first order in η_T

$$\begin{aligned}\mathcal{M} &= D_g^{(r)} - 4n^i\partial_i V + \frac{1}{2}n_{ij}M^{ij} - \eta_T[\dot{D}_g^{(r)} - 4n^i\partial_i\dot{V} + \frac{1}{2}n_{ij}\dot{M}^{ij} \\ &\quad + n^j\partial_j D_g^{(r)} - 4n^in^j\partial_i\partial_j V + \frac{1}{2}n^in_{kj}\partial_i M^{kj} - 4n^j\partial_j(\Phi - \Psi)]\ . \qquad (9.72)\end{aligned}$$

Using (9.72) to calculate the drag force yields

$$F_i = (\rho_r/3)[4\partial_i\dot{V} - \partial_i D_g^{(r)} + 4\partial_i(\Phi - \Psi)]\ .$$

Inserting F_i in (9.70), we obtain

$$(\rho_b + (4/3)\rho_r)\partial_i\dot{V} + \rho_b(\dot{a}/a)\partial_i V = (\rho_r/3)\partial_i D_g^{(r)} + (\rho_b + (4/3)\rho_r)\partial_i\Psi - (4\rho_r/3)\partial_i\Phi\ .$$

This is equivalent to momentum conservation, the second equation of (7.55) for $\rho = \rho_b + \rho_r$, $p = \rho_r/3$ and $\Gamma = \Pi = 0$, if we use

$$D_g^{(r)} = (4/3)D_g^{(b)} \quad \text{and} \quad D_g = \frac{\rho_r D_g^{(r)} + \rho_b D_g^{(b)}}{\rho_b + \rho_r}\ .$$

In this limit therefore, baryons and photons behave like a single fluid with density $\rho = \rho_r + \rho_b$ and pressure $p = \rho_r/3$.

From (7.55) we can derive a second order equation for D_g. This equation can be simplified if expressed in terms of the variable D related by (8.1). To discuss the coupled matter radiation fluid we consider a plane wave $D = D(t)\exp(i\boldsymbol{k}\cdot\boldsymbol{x})$. We then obtain

$$\ddot{D} + c_s^2k^2D + (1 + 3c_s^2 - 6w)(\dot{a}/a)\dot{D} - 3[w(\ddot{a}/a) - (\dot{a}/a)^2(3(c_s^2 - w) - (1/2)(1+w))]D = 0\ .$$

For small wavelengths (sub-horizon), which are however sufficiently large for the fluid approximation to be valid, $1/\eta_T \gg c_sk \gg 1/\eta$, we may drop the term in square brackets. The ansatz $D(t) = A(t)\exp(-i\int kc_sdt)$ then eliminates the term of order $c_s^2k^2$. For the terms of order c_sk/η we obtain the equation

$$2\dot{A}/A + (1 - 3c_s^2 - 6w)(\dot{a}/a) + \dot{c}_s/c_s = 0\ . \qquad (9.73)$$

For the case $c_s^2 = w$ =const. , this equation is solved by $A \propto (k\eta)^{1-\nu}$ with $\nu = 2/(3w+1)$, i.e., the short wave limit. In our situation we have

$$\begin{aligned}w &= \frac{\rho_r}{3(\rho_r + \rho_b)} \\ c_s^2 &= \frac{\dot{\rho}_r}{3(\dot{\rho}_r + \dot{\rho}_b)} = \frac{(4/3)\rho_r}{4\rho_r + 3\rho_b} \quad \text{and} \\ \dot{c}_s/c_s &= -3/2(\dot{a}/a)\frac{\rho_b}{4\rho_r + 3\rho_b}\ .\end{aligned}$$

Using all this, one finds that

$$A = \left(\frac{\rho_b + (4/3)\rho_r}{c_s(\rho_r + \rho_b)^2a^4}\right)^{1/2} = \left(\frac{\rho + p}{c_s\rho^2a^4}\right)^{1/2}$$

solves (9.73) exactly, so that we finally obtain the approximate solution for the, tightly coupled matter radiation fluid on sub-horizon scales

$$D(t) \propto \left(\frac{\rho + p}{c_s \rho^2 a^4}\right)^{1/2} \exp(-ik \int c_s d\eta) \ . \tag{9.74}$$

Note that this short wave approximation is only valid in the limit $\eta \gg 1/(c_s k)$, thus the limit to the matter dominated regime, $c_s \to 0$ cannot be performed. In the limit to the radiation dominated regime, $c_s^2 \to 1/3$ and $\rho \propto a^{-4}$ we recover the acoustic waves with constant amplitude which we have already found in the last subsection. But also in this limit, we still need matter to ensure $\eta_T = 1/(a\sigma_T n_e) \ll \eta$. In the opposite case, $\eta_T \gg \eta$, radiation does not behave like an ideal fluid but it becomes collisionless and has to be treated with Liouville's equation ((9.64) without the collision term).

9.3.3 Damping by photon diffusion

In this subsection we discuss the Boltzmann equation in the next order, $(\eta_T/\eta)^2$. In this order we will obtain the damping of fluctuations in an ionized plasma due to the finiteness of the mean free path; the non-perfect coupling. We follow the treatment by Peebles [1980] [65] (using our gauge-invariant approach instead of synchronous gauge). Again we consider Eqs. (9.64) and (9.70), but since we are mainly interested in collisions which take place on time scales $\eta_T \ll \eta$, we neglect gravitational effects and the time dependence of the coefficients. We can then look for solutions of the form

$$V \propto \mathcal{M} \propto \exp(i(\boldsymbol{k} \cdot \boldsymbol{x} - \omega\eta)) \ .$$

In (9.64) and (9.70) this yields (neglecting also the angular dependence of Compton scattering, i.e., the term $n_{ij}M^{ij}$)

$$\mathcal{M} = \frac{1}{4}\frac{D_g^{(r)} - ik\boldsymbol{n}V}{1 - i\eta_T(\omega - \boldsymbol{k}\cdot\boldsymbol{n})} \tag{9.75}$$

and

$$\eta_T \boldsymbol{k}\omega V = (4\rho_r/3\rho_b)(i\boldsymbol{k}V + \boldsymbol{M}) \ , \tag{9.76}$$

with $\boldsymbol{M} = (3/4\pi)\int \boldsymbol{n}\mathcal{M}d\Omega$. Integrating (9.64) over angles, one obtains $\dot{D}_g^{(r)} + (1/3)\partial_i M^i = 0$. With our ansatz therefore $\boldsymbol{k}\cdot\boldsymbol{M} = 3\omega D_g^{(r)}$. Using this after scalar multiplication of (9.76) with $\boldsymbol{k}$, we find, setting $R = 3\rho_b/4\rho_r$,

$$V = \frac{(3/4)\omega D_g^{(r)}}{\eta_T k^2 R\omega - ik^2} \ .$$

Inserting this result for V in (9.75) leads to

$$\mathcal{M} = \frac{D_g^{(r)}}{4}\frac{1 + \frac{3\mu\omega/k}{1 - i\eta_T\omega R}}{1 - i\eta_T(\omega - k\mu)} \ ,$$

where we have set $\mu = \hat{\mathrm{k}}\cdot\mathrm{n}$. This is the result of Peebles [1980] [65], where this calculation is performed in synchronous gauge. Like in there (§92), one obtains in lowest order $\omega\eta_T$ the dispersion relation. Using

$$\frac{1}{2}\int_{-1}^{1}\mathcal{M}d\mu = \frac{D_g^{(r)}}{4} \ , \quad \text{which yields} \quad 1 = \frac{1}{2}\int_{-1}^{1}\frac{1 + \frac{3\mu\omega/k}{1 - i\eta_T\omega R}}{1 - i\eta_T(\omega - k\mu)}d\mu$$

one finds

$$\omega = \omega_0 - i\gamma \quad \text{with} \quad \omega_0 = k/[3(1+R)]^{1/2} \quad \text{and} \quad \gamma = (k^2\eta_T/6)\frac{R^2 + \frac{4}{5}(R+1)}{(R+1)^2} \ . \tag{9.77}$$

In the baryon dominated regime, $R \geq 1$, therefore

$$\gamma \approx k^2 \eta_T / 6 \ . \tag{9.78}$$

(If the angular dependence of Thompson scattering is not neglected, the term $\frac{4}{5}(R+1)$ becomes $\frac{8}{9}(R+1)$. If also polarization is taken into account, one obtains $\frac{16}{15}(R+1)$.)

Posing $k_{\rm damp}\eta_T/6 = 1$, this leads to a damping scale $\lambda_{\rm damp} \sim \eta_T(\eta_{\rm dec})$, which is projected in the microwave sky to an angle

$$\vartheta_{\rm damp} \sim \frac{\eta_T(\eta_{\rm dec})}{\chi(\eta_0 - \eta_{\rm dec})} \ .$$

For $\kappa = 0$ this corresponds to a few arc minutes and to the harmonic number

$$\ell_{\rm damp} = \pi/\vartheta_{\rm damp} \simeq \frac{\pi\eta_0}{10\eta_T(\eta_{\rm dec})} \simeq \frac{(1+z_{\rm dec})^2}{10}\Omega_b h \ . \tag{9.79}$$

This estimate is very crude since we are using the the approximation for η_T from the tight coupling regime just where coupling stops to be tight, and we assume an arbitrary value of $n_e \sim 0.1 n_B$ at the moment of decoupling. Both these effects lower $\ell_{\rm damp}$ somewhat and numerical results give

$$\ell_{\rm damp} \sim 600$$

in a $\kappa = 0$ universe. In open (closed) universes, this scale (which of course also depends on Ω_b) is moved to larger (lower) ℓ due to projection. A reasonable approximation for the damping harmonic is

$$\ell_{\rm damp} \sim 600 \left(\frac{\Omega_b h}{0.01\Omega^{1/2}}\right) \ .$$

Temperature fluctuations on smaller scales, $\ell > \ell_{\rm damp}$ are exponentially damped by photon diffusion.

9.4 Polarization

Thomson scattering is not isotropic. And what is more, for a non–isotropic photon distribution it depends on the polarisation: Even if the incident photon beam is unpolarised, the scattered beam will be, unless the incident distribution is perfectly isotropic. In the previous section we have neglected this effect by summing over the initial polarizations and averaging over final polarizations. Now we want to derive the difference in the Boltzmann equation taking into account polarisation. For simplicity (and since this is by far the most relevant case) we concentrate on scalar perturbations.

Polarisation is usually characterized by means of the *Stokes parameters* [66, 67, 68].

For a harmonic electromagnetic wave with associated electric field

$$\mathbf{E}(\mathbf{x},t) = (\boldsymbol{\varepsilon}_1 E_1 + \boldsymbol{\varepsilon}_2 E_2)\, e^{i\omega(\mathbf{nx}-t)} \ , \tag{9.80}$$

where $\mathbf{n}$, $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ form an orthonormal basis and the complex field amplitudes are parameterised as $E_j = a_j e^{i\delta_j}$, the Stokes parameters are given by

$$I = a_1^2 + a_2^2 \tag{9.81}$$
$$Q = a_1^2 - a_2^2 \tag{9.82}$$
$$U = 2a_1a_2\cos(\delta_2 - \delta_1) \tag{9.83}$$
$$V = 2a_1a_2\sin(\delta_2 - \delta_1). \tag{9.84}$$

I is the intensity of the wave (whose perturbation $\imath$ has been introduced in the previous section), while Q is a measure of the strength of linear polarisation (the ratio of the principal axis of the polarisation ellipse). U and V give phase information (the orientation of the ellipse). One can show that U and V are not coupled to I and Q by Thomson scattering. For scalar perturbations U and V even vanish. We therefore ignore them here.

Since Q vanishes in the background, to first order it obeys the unperturbed Liouville equation,

$$\partial_\eta \mathcal{M}^{(Q)} + in^\ell k_\ell \mathcal{M}^{(Q)} - \Gamma^{(S)\,i}_{jk} n^j n^k \frac{\partial \mathcal{M}^{(Q)}}{\partial n^i} = 0. \tag{9.85}$$

The differential cross section of Thomson scattering for a photon with incident polarisation $\varepsilon_{(i)}$ and outgoing polarisation $\varepsilon_{(s)}$ is

$$\frac{d\sigma}{d\Omega} = \frac{3}{8\pi}\sigma_T \left|\varepsilon^*_{(s)}\varepsilon_{(i)}\right|^2 . \tag{9.86}$$

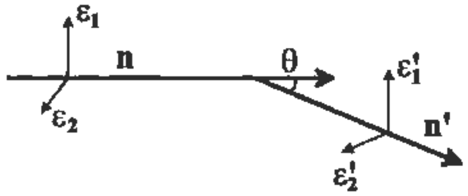

Figure 9.4: Definition of the angles and vectors for Thomson scattering in the $(\mathbf{n}, \varepsilon_2)$ plane.

It is often convenient to introduce the two 'partial' intensities $I_1 \equiv a_1^2 = (I+Q)/2$ and $I_2 \equiv a_2^2 = (I-Q)/2$. A wave scattered in the $(\mathbf{n}, \varepsilon_2)$ plane (see Fig. 9.4) by an angle θ has the intensities

$$\begin{aligned} I_1^{(s)} &= \frac{3\sigma_T}{8\pi} I_1^{(i)} \\ I_2^{(s)} &= \frac{3\sigma_T}{8\pi} I_2^{(i)} \cos^2\theta, \end{aligned} \tag{9.87}$$

or, expressed in terms of the Stokes parameters,

$$\begin{pmatrix} I^{(s)} \\ Q^{(s)} \end{pmatrix} = \frac{3\sigma_T}{16\pi} \begin{pmatrix} 1+\cos^2\theta & \sin^2\theta \\ \sin^2\theta & 1+\cos^2\theta \end{pmatrix} \begin{pmatrix} I^{(i)} \\ Q^{(i)} \end{pmatrix}. \tag{9.88}$$

To rotate a wave into a common coordinate system, one uses that a rotation in the $(\varepsilon_1, \varepsilon_2)$ plane by an angle ϕ. According to the general transformation of the Stokes parameters under rotation, this brings (I, Q) into (I', Q') given by

$$I' = I, \quad Q' = Q\cos(2\phi), \quad \text{or} \tag{9.89}$$

$$\begin{pmatrix} I_1' \\ I_2' \end{pmatrix} = \begin{pmatrix} \cos^2\phi & \sin^2\phi \\ \sin^2\phi & \cos^2\phi \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix}. \tag{9.90}$$

If we start with a wave $(I^{(i)}, Q^{(i)})$ propagating in the direction $\mathbf{n}$ that is scattered into a wave $(I^{(s)}, Q^{(s)})$ in direction $\mathbf{n}'$, then we need to go through the following steps (we will use the plane (z, y) as reference plane, see Fig. (9.5) for definitions of the angles and vectors):

1. Rotate the plane $(\mathbf{n}, \mathbf{n}')$ around $\mathbf{n}$ into the plane $(z, \mathbf{n})$. One needs to apply the rotation (9.89) for $\phi = \alpha$ to the Stokes parameters.

2. Rotate the new plane $(\mathbf{n}, \mathbf{n}')$ around z into the reference plane. This operation does not influence the scattering.

3. Now we are in the known case of (9.87) and (9.88). Hence we can apply one of those scattering matrices.

4. We rotate the scattering plane back around z into the old $(z, \mathbf{n}')$ plane. This does again not change the Stokes parameters.

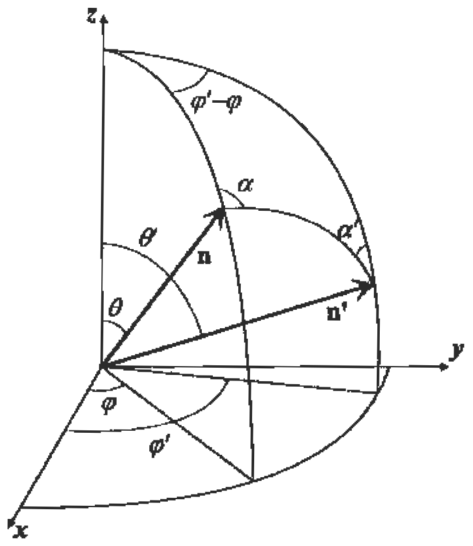

Figure 9.5: Definition of the angles and vectors for Thomson scattering in the general case. The polarisation vectors are oriented like in Fig. 9.4.

5. Finally we rotate around $\mathbf{n}'$ by the angle α' to reach the original state. To do this, we have to apply the rotation matrix (9.89) again, but for $\phi = \alpha'$.

The scattering matrix for an incoming photon with direction $\mathbf{n}$ that is scattered into direction $\mathbf{n}'$ for a reference frame with $\hat{\mathbf{k}} = z$ and for the scalar part is then given by:

$$\begin{pmatrix} I^{(s)} \\ Q^{(s)} \end{pmatrix} = \frac{3\sigma_T}{16\pi} \begin{pmatrix} 1 + (\mathbf{n}\cdot\mathbf{n}')^2 & (\mathbf{n}\cdot\mathbf{n}')^2 - (\mathbf{n}\cdot\hat{\mathbf{k}})^2 \\ (\mathbf{n}\cdot\mathbf{n}')^2 - (\mathbf{n}'\cdot\hat{\mathbf{k}})^2 & 1 + (\mathbf{n}\cdot\mathbf{n}')^2 - (\mathbf{n}\cdot\hat{\mathbf{k}})^2 - (\mathbf{n}'\cdot\hat{\mathbf{k}})^2 \end{pmatrix} \begin{pmatrix} I^{(i)} \\ Q^{(i)} \end{pmatrix} \tag{9.91}$$

To calculate the collision term including polarisation effects, we change into the (I_1, I_2) basis . For each of the two intensities $\lambda \in \{1,2\}$ we then have the collision term given by

$$C[f^{(\lambda)}] = \frac{df_+^{(\lambda)}}{d\eta} - \frac{df_-^{(\lambda)}}{d\eta}, \tag{9.92}$$

where $f_+^{(\lambda)}$ and $f_-^{(\lambda)}$ denote the distribution of photons in the polarisation state λ scattered into respectively out of the beam due to Compton scattering.

In the matter (baryon/electron) rest frame, which we indicate by a prime, we know that

$$\frac{df_+^{(\lambda)\prime}}{dt'}(p, \mathbf{n}) = \frac{\sigma_T n_e}{4\pi} \int f^{(\delta)\prime}(p', \mathbf{n}')\omega_\delta^\lambda(\mathbf{n}, \mathbf{n}')d\Omega' ,$$

where n_e denotes the electron number density, σ_T is the Thomson cross section, and ω_δ^λ is the normalised Thomson scattering matrix (9.91), but in the basis $\{I_1, I_2\}$.

Using the Lorentz transformation from the baryon rest frame to the laboratory frame (like in the previous section) and performing the integration over photon energies, we obtain

$$\rho_\gamma \frac{d\iota_+^{(\lambda)}(\mathbf{n})}{dt'} = \rho_\gamma \sigma_T n_e \left[1 + 4n_i(v^i - B^i) + \frac{1}{4\pi} \int \iota^{(\delta)}(\mathbf{n}')\omega_\delta^\lambda(\mathbf{n}, \mathbf{n}')d\Omega' \right] .$$

The distribution of photons scattered out of the beam is like in the previous section,

$$\frac{df_-^{(\lambda)}}{dt'} = \sigma_T n_e f^{(\lambda)\prime}(p', \mathbf{n}) ,$$

so that we finally have

$$\begin{aligned} C^{(\lambda)\prime} &= \frac{4\pi}{\rho_\gamma a^4}\int dp(\frac{df_+^{(\lambda)}}{dt'} - \frac{df_-^{(\lambda)}}{dt'})p^3 \\ &= \frac{1}{2}\sigma_T n_e\left[-\iota^{(\lambda)} + 4n_i(v^i - B^i) + \frac{1}{4\pi}\int \iota^{(\delta)}(\mathbf{n}')\omega_\delta^\lambda(\mathbf{n},\mathbf{n}')d\Omega'\right] . \end{aligned}$$

We convert this result to the normal Stokes parameters by setting $C^{(I)} = C^{(1)} + C^{(2)}$ and $C^{(Q)} = C^{(1)} - C^{(2)}$ as well as $\iota = \iota^{(1)} + \iota^{(2)}$ and $q = \iota^{(1)} - \iota^{(2)}$. The resulting collision integrals are then

$$C^{(I)\prime} = \sigma_T n_e\left[-\iota + 4n_i(v^i - B^i) + \frac{1}{4\pi}\int(\tilde\omega_{11}\iota + \tilde\omega_{12}q)\,d\Omega'\right] \tag{9.93}$$

$$C^{(Q)\prime} = \sigma_T n_e\left[-q + \frac{1}{4\pi}\int(\tilde\omega_{21}\iota + \tilde\omega_{22}q)\,d\Omega'\right] \tag{9.94}$$

where $\tilde\omega$ is the normalised scattering matrix for I and Q from Eq. (9.91). Clearly, $q = \mathcal{M}^{(Q)}$. The term $\tilde\omega_{11}$ is as in the previous section,

$$\tilde\omega_{11}(\mathbf{n},\mathbf{n}') = 3/4[1 + (\mathbf{nn}')^2] = 1 + \frac{3}{4}n_{ij}n'_{ij} \quad \text{with} \quad n_{ij} \equiv n_i n_j - \frac{1}{3}\delta_{ij} . \tag{9.95}$$

Using $\delta_\gamma = (1/4\pi)\int\iota(n)d\Omega$ (the photon energy density perturbation) and the definitions of the gauge-invariant variables $\mathcal{M}^{(S)}$ and $V^{(b)}$ for the photon brightness perturbation and the baryon velocity potential, we can write $C^{(I)\prime}$ in gauge-invariant form.

$$C^{(I)\prime} = \sigma_T n_e\left[D_g^{(r)} - \mathcal{M}^{(S)} + 4n_i\partial_i V^{(b)} + \frac{1}{2}n_{ij}M^{ij} + \frac{1}{4\pi}\int\tilde\omega_{12}\,q\,d\Omega'\right] , \tag{9.96}$$

The scattering matrix element $\tilde\omega_{12} = {}^3/_4[(\mathbf{nn}')^2 - (\mathbf{n}\hat{\mathbf{k}})^2]$ can be rewritten as ${}^3/_4[(\mathbf{nn}')^2 - {}^1/_3] - {}^3/_4[(\mathbf{n}\hat{\mathbf{k}})^2 - {}^1/_3]$. The first part then gives ${}^1/_2 n_{ij}M^{(Q)ij}$ just like for the brightness perturbation.

Since the term in square brackets of (9.94) and (9.96) is already first order we can set $dt' = dt$ which yields $C = \frac{dt'}{d\eta}C' = \frac{dt}{d\eta}C' = aC'$. The Boltzmann equation for scalar perturbations expressed in terms of the gauge invariant variable $\mathcal{M}^{(S)}$ then becomes

$$\begin{aligned} \dot{\mathcal{M}}^{(S)} + n^i\partial_i\mathcal{M}^{(S)} &= 4n^i\partial_i(\Phi - \Psi) + a\sigma_T n_e\Big[D_g^{(\gamma)} - \mathcal{M}^{(S)} - 4n^i\partial_i V^{(b)} \\ &+\frac{1}{2}n_{ij}\left(M^{ij} + M^{(Q)ij}\right) - \frac{3}{4}\left((\mathbf{n}\hat{\mathbf{k}})^2 - \frac{1}{3}\right)\frac{1}{4\pi}\int\mathcal{M}^{(Q)}(\mathbf{n}')d\Omega'\Big] . \end{aligned} \tag{9.97}$$

Note the difference to the result obtained neglecting polarisation (Eq. 9.49)!

For the polarisation equation, we rewrite the other two matrix elements correspondingly. We find then

$$\begin{aligned} &\dot{\mathcal{M}}^{(Q)} + n^i\partial_i\mathcal{M}^{(Q)} = \\ &a\sigma_T n_e\Big[-\mathcal{M}^{(Q)} + \frac{1}{2}n_{ij}\left(M^{ij} + M^{(Q)ij}\right) \\ &-\frac{3}{4}\frac{1}{4\pi}\int(\mu'^2 - 1/3)\left(\mathcal{M}^{(S)}(\mathbf{n}') + \mathcal{M}^{(Q)}(\mathbf{n}')\right)d\Omega' \\ &-\frac{3}{4}\left((\mathbf{n}\hat{\mathbf{k}})^2 - 1/3\right)\frac{1}{4\pi}\int\mathcal{M}^{(Q)}(\mathbf{n}')d\Omega' + \frac{1}{2}\frac{1}{4\pi}\int\mathcal{M}^{(Q)}(\mathbf{n}')d\Omega'\Big] . \end{aligned} \tag{9.98}$$

The $C_\ell^{(P)}$'s for the polarisation are now obtained from $\mathcal{M}^{(Q)} = \left(\frac{\Delta T}{T}\right)^{(Q)}$ in exactly the same way as the once for the temperature anisotropies from $\left(\frac{\Delta T}{T}\right)^{(S)}$ by Eqn. (9.22). The figure below shows both, the temperature and polarisation C_ℓ's as well as the cross correlation, $\langle\mathcal{M}^{(S)}(\mathcal{M}^{(Q)})^*\rangle$.

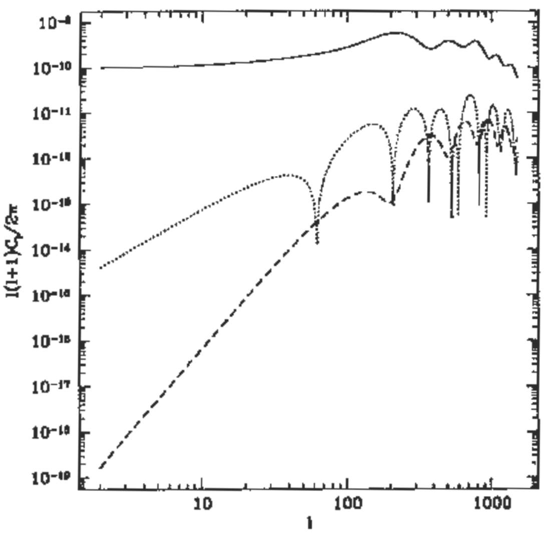

Figure 9.6: The temperature anisotropy (solid), the polarisation (dashed) and their correlation (dotted) are shown for a purely scalar standard CMD model.

9.5 Summary

9.5.1 Physical processes

- Before recombination, photons and baryons form a tightly coupled fluid which performs acoustic oscillations on sub-horizon scales.
- Depending on the initial conditions, these oscillations are sine waves (isocurvature case) or cosine waves (adiabatic case).
- After recombination, the photons move along perturbed geodesics, only influenced by the metric perturbations.
- Vector perturbations of the metric decay as a^{-2} after creation and their effects on CMB anisotropies are negligible for models where initial fluctuations are created early, *e.g.* during an inflationary phase. This is different for models which constantly seed fluctuations in the geometry, *e.g.* topological defects.
- Tensor perturbations of the metric have constant amplitude on super-horizon scales and perform damped oscillations $\propto a^{-1}$ once they enter the horizon.
- Scalar perturbations of the metric are roughly constant if they enter the horizon only after the time of matter and radiation equality. On scales which enter the horizon before equality they are damped by a factor $(z_{eq}/z_{in})^2$, where z_{eq} and z_{in} are the redshift of equality and of horizon crossing, respectively.
- Perturbations on small scales, $k \gtrsim k_T \simeq (\Omega_b h/20)(z_{dec}+1)^2 H_0$ are exponentially damped by collisional damping during recombination (Silk damping).

9.5.2 Scale dependence

- On large scales (larger than the horizon scale at recombination, $\ell \lesssim \ell_H \simeq \pi/\vartheta_H$, with $\vartheta_H = \eta_{dec}/\chi(\eta_0 - \eta_{dec})$, perturbations are dominated by gravitational effects: Inflationary models typically lead to $k^3\left\langle|\Psi-\Phi|^2(k,\eta_{\rm dec})\right\rangle \simeq$ const. and $k^3\left\langle H^2\right\rangle \simeq$ const. on these scales. This implies a flat "Harrison-Zel'dovich" spectrum,

$$\left(\frac{\Delta T}{T}\right)^2(\vartheta_\ell) \simeq \ell(\ell+1)C_\ell \simeq \text{const.}, \quad \vartheta_\ell = \frac{\pi}{\ell}. \tag{9.99}$$

- On intermediate scales, $\ell_H < \ell < \ell_{\rm damp} \sim 600$, CMB anisotropies mainly reflect the acoustic oscillations of the photon/baryon plasma prior to recombination. The position of the first peak is severely affected by initial conditions (adiabatic or isocurvature). For $\kappa = 0$, the first contraction peak is at about $\ell_1^{(a)} \sim 220$ if the initial conditions are adiabatic, while the first contraction peak is at $\ell_1^{(i)} \sim 350$ for isocurvature initial conditions. The amplitude of and the distance between the peaks depend strongly on cosmological parameters.
- On small scales, $\ell_{\rm damp} < \ell$, fluctuations are collisionally damped during recombination ("Silk damping"). The damping scale depends mainly on $\Omega_B h$ and Ω.

9.5.3 The main influence of cosmological parameters

- **Curvature, $h^2\Omega_\kappa$:**

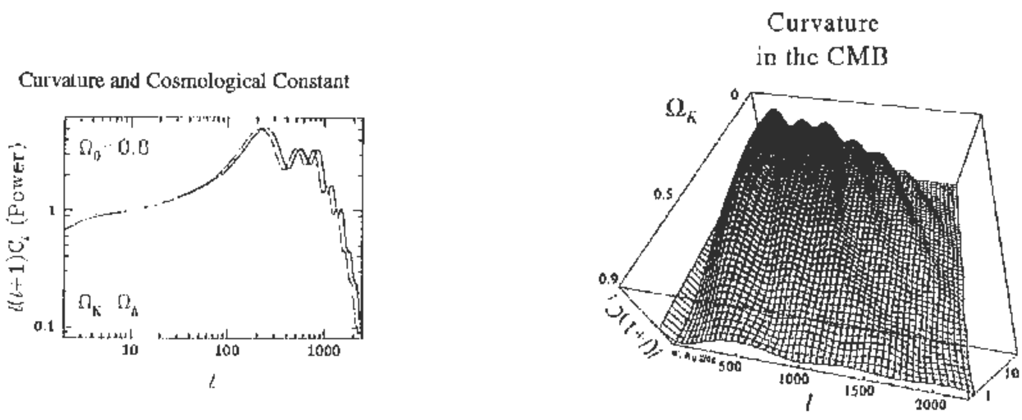

Figure 9.7: The temperature anisotropy $\ell(\ell+1)C_\ell$'s are shown as a function of Ω_κ (only the open case, $\Omega_\kappa \geq 0$ is considered). The top panel shows the difference between the action of curvature and of a cosmological constant for fixed $\Omega_m = 0.8$ (Taken from Wayne Hu's homepage, http://www.sns.ias.edu/~whu).

  - Mainly affects the inter-peak distance, $\Delta\ell$, and, for given initial conditions, the position of the first peak. Positive curvature lowers $\Delta\ell$ while negative curvature enhances it (see Fig. 9.7).
  - Curvature also leads to an integrated Sachs–Wolfe contribution which is especially important for $\kappa > 0$ at very low ℓ. Overall, this leads to some enhancement of the Sachs–Wolfe contribution and therefore (after normalisation to the COBE measurements) to somewhat lower acoustic peaks.
  - Non-zero curvature changes the epoch of equal matter and radiation, leading to an enhancement of the acoustic peaks if $\kappa > 0$ and to a decrease if $\kappa < 0$ (see Fig. 9.7).

- **Baryon density, $\rho_B = \Omega_B h^2 \cdot 10^{-29}\text{g/cm}^3$:**

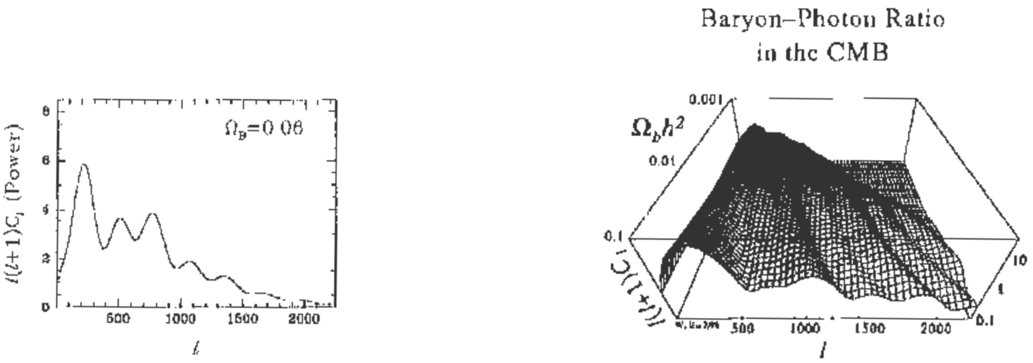

Figure 9.8: The temperature anisotropy $\ell(\ell+1)C_\ell$'s are shown as a function of the baryon density, $\Omega_B h^2$. (Taken from Wayne Hu's homepage, http://www.sns.ias.edu/~whu).

- A high baryonic density enhances the compression peaks and decreases the expansion peaks via the self–gravity of the baryons.
- It also reduces the damping scale, $\lambda_T = 1/(a_{\rm dec}\sigma_T n_e(\eta_{\rm dec}))$, leading to an increase in $\ell_{\rm damp}$.
- By its influence on the plasma sound velocity, $c_s = 1/3(1 + \dot\rho_B/\dot\rho_\gamma)^{-1}$, it prolongs the oscillation period (cf Fig. 9.8).

• **Cosmological Constant,** $\Lambda = \frac{\Omega_\Lambda h^2}{8\pi G} \cdot 10^{-29}{\rm g/cm}^3$**:**

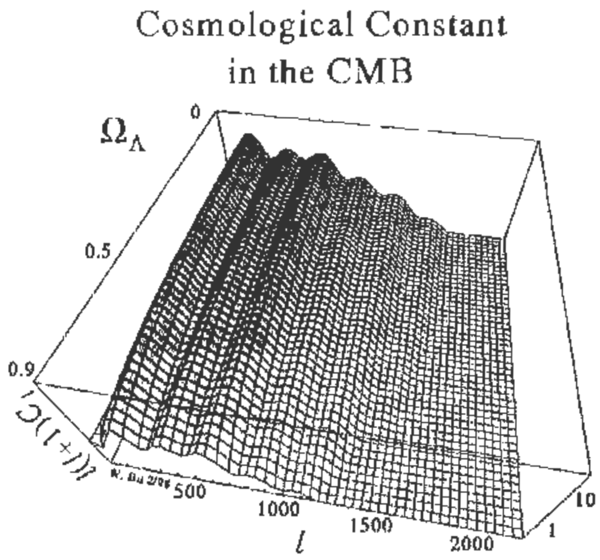

Figure 9.9: The temperature anisotropy $\ell(\ell+1)C_\ell$'s are shown as a function of the cosmological constant, Ω_Λ. (Taken from Wayne Hu's homepage, http://www.sns.ias.edu/~whu).

The presence of a cosmological constant at fixed $\Omega_{\rm tot} = \Omega_m + \Omega_\Lambda$ delays the epoch of equal matter and radiation. During the radiation dominated era, the gravitational potential is not constant, but decays as soon as a given scale enters the horizon. If $\eta_{\rm eq} \sim \eta_{\rm dec}$ this induces an integrated Sachs–Wolfe (ISW) contribution which boosts mainly the first acoustic peak. If Ω_Λ becomes very large $\gtrsim 0.8$ it also boosts the late integrated Sachs–Wolfe contribution and the relative height of the acoustic peaks begins to decrease again (see Fig. 9.9).

- **Hubble Parameter, $H_0 = 100h$ km/(s Mpc):**

  For fixed curvature and cosmological constant, lowering the Hubble parameter also delays the epoch of equal matter and radiation, $\eta_{\rm eq} \to \eta_{\rm dec}$, since

$$z_{\rm eq} + 1 = \frac{\Omega_m}{\Omega_{\rm rad}} \simeq 2.4 \cdot 10^4 \Omega_m h^2. \tag{9.100}$$

  Therefore the same type of ISW contribution as for Λ–models boosts the first acoustic peak.

- **Initial conditions:**
  - A tensor contribution enhances the large scales fluctuations but not the acoustic peaks, thereby lowering their relative amplitude.
  - A "blue" fluctuation spectrum, $n > 1$, enhances fluctuations on smaller scales and raises thereby the acoustic peaks.

Chapter 10

Observations

In this short, final chapter we want to discuss briefly the experimental situation. It has been clear for a long time that, if initial fluctuations have led to the formation of large scale structure by gravitational instability, they should have induced fluctuations in the cosmic microwave background [71, 72]. Until spring 1992, however, only the dipole anisotropy had been detected [61, 62]. Its value is [56]

$$\left\langle\left(\frac{\Delta T}{T}\right)^2\right\rangle^{\text{dipole}} = (1.528 \pm 0.004) \times 10^{-6} \ .$$

After many upper limits, the DMR experiment aboard the COBE satellite measured for the first time convincingly positive anisotropies [57]. It found

$$\left\langle\left(\frac{\Delta T}{T}\right)^2\right\rangle(\theta) \sim (30\mu\text{K})^2 \tag{10.1}$$

on all angular scales $\theta \geq 7°$. Many more positive measurements have been performed since then. A complete list until September 1999 is given in the Table 10.1 and indicated in Fig. 10.1.

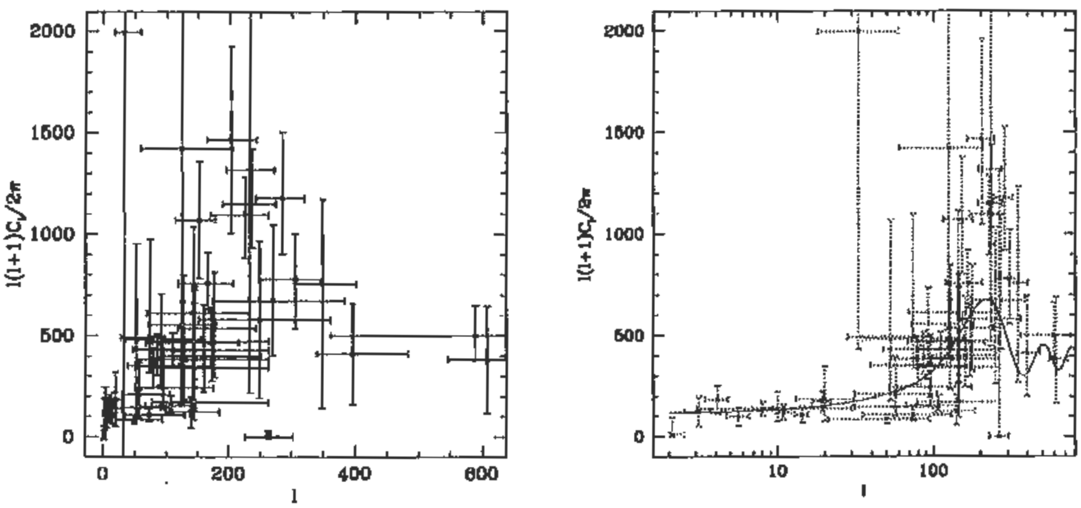

Figure 10.1: The measured temperature anisotropies, $\ell(\ell+1)C_\ell$ indicated in the table above are shown in a lin-lin plot (left) and in a log-lin plot with the theoretical curve from a standard, adiabatic cold dark matter model (right).

As one sees in the above figure, present data, apart from COBE, is very scattered. It may well be that many of these experiments still have normalization problems which are more severe than

Experiment	effective ℓ	$\Delta T^2(\mu K)^2$	$+(\mu K)^2$	$-(\mu K)^2$	Sky Coverage	Reference
COBE1	2.1	11.5	84.0	11.5	0.65	[82]
COBE2	3.1	125	74.7	75.5	0.65	[82]
COBE3	4.1	184	69.4	69.7	0.65	[82]
COBE4	5.6	100	45.8	45.8	0.65	[82]
COBE5	8.0	137.6	35.8	35.7	0.65	[82]
COBE6	10.9	122	36.8	36.5	0.65	[82]
COBE7	14.3	108	39.6	39.6	0.65	[82]
COBE8	19.4	173	51	52	0.65	[82]
FIRS	10	137.6	82.7	62.6	-	[83]
Tenerife	20.1	185	160.5	110.8	0.0124	[84]
IAC/Bartol	33.0	1996	2989	1566	-	[85]
IAC/Bartol	53.0	481	588	308	-	[85]
PYTHONV	50	84	23.4	20.5	0.01	[86]
PYTHONV	74	107	35.7	30.	0.01	[86]
PYTHONV	108	153	53.3	37	0.01	[86]
PYTHONV	140	125	81.5	67.3	0.01	[86]
PYTHONV	172	464	188	170	0.01	[86]
PYTHONV	203	1467	494	423	0.01	[86]
PYTHONV	233	1318	1090	871	0.01	[86]
BAM	74	492	604	158	-	[87]
QMAP-F1+2Ka	80	351	95	97	0.01	[88]
QMAP-F1+2Ka	126	554	118	124	0.01	[88]
QMAP-Q	111	430	87	79	0.01	[88]
South Pole 91	57	145	98	48	0.005	[89]
South Pole 94	57	210	187	65	0.005	[89]
PYTHON	92	464	271	183	-	[90]
PYTHON	177	535	313	213	-	[90]
ARGO Hercules	95	243	120	96	0.0024	[101]
ARGO Aries	95	349	156	157	0.0024	[91]
Saskatoon	96	382	125	78	0.0037	[92]
Saskatoon	166	758	161	126	0.0037	[92]
Saskatoon	236	1150	286	206	0.0037	[92]
Saskatoon	285	1177	518	258	0.0037	[92]
Saskatoon	348	758	475	490	0.0037	[92]
IAM	125	1421	1535	979	-	[93]
TOCO	128	481	367	252	-	[94]
TOCO	152	1070	306	268	-	[94]
TOCO	226	1096	193	201	-	[94]
TOCO	780	223	245	247	-	[94]
MSAM94	143	612	503	353	0.0007	[95]
MSAM94	248	581	351	322	0.0007	[95]
MSAM95	160	398	295	156	0.0007	[96]
MSAM95	270	672	424	242	0.0007	[96]
MAX HR	145	170	132	75	0.0002	[97]
MAX PH	145	473	373	161	0.0002	[97]
MAX GUM	145	473	227	170	0.0002	[97]
MAX ID	145	341	397	171	0.0002	[97]
MAX SH	145	384	416	214	0.0002	[97]
CAT1	396	411	287	211	0.0001	[99]
CAT2	608	382	309	218	0.0001	[99]
RING5m	589	499	163	111	-	[100]

Table 10.1: The published CMB anisotropy detections until September 1999. The 3., 4. and 5. column denote the value of the anisotropy and the upper and lower 1-σ errors respectively.

indicated in the error bars. So far only one experiment sees a rather well distinguished 'raise and fall' which may be due to the first acoustic peak. All the other experiments on small scales only indicate an acoustic peak when compared to COBE. This situation will change drastically once the data from the second BOOMERanG flight will be analyzed or when the MAP data arrives (see below).

The experiments can be split into three classes: Satellite experiments, balloon–borne experiments and ground based experiments. The technical and economical advantages of ground based experiments are obvious. Their main problem is atmospheric fluctuation. This can be reduced by two methods:

- Choose a very high altitude and very cold site, *e.g.* the south pole. Several experiments like SP, Python and White Dish have chosen this site.
- Measure anisotropies on very small scales, preferably by interferometry (CAT, VSA, Jodrell Bank).

Balloon–borne experiments flying at about 40km altitude have less problems with the Earths atmosphere but they face the following difficulties:

- They are limited in weight.
- They cannot be manipulated at will in flight.
- They have a rather short duration.
- To secure all the data taken on the balloon, they have to be recovered intact.

Yet the advantages of overcoming the atmosphere are so significant that many groups have chosen this approach, like *e.g.* MAX, FIRS, MSAM, QMAP, TopHat, etc. The BOOMERanG experiment even combined the two advantages and performed a long–duration flight (10 days) on the south pole in December 1998. (Unfortunately it will still take a considerable amount of time until the data will be fully analyzed. But preliminary maps look very promising and indicate that this dataset is an entirely different quality than everything we have so far. The BOOMERanG-98 data should reveal the C_ℓs from $\ell \sim 60$ to $\ell \sim 600$ with about 15% errors!)

The third possibility are satellite experiments. They avoid atmospheric problems altogether, but this solution is very expensive. So far only one satellite has been launched (namely COBE in 1989) and two more are planned: MAP (Microwave Anisotropy Probe, a NASA MIDEX mission, to be launched in 2001) and PLANCK, an ESA medium size mission of the "Horizon 2000" program, to be launched in 2006.

MAP will perform measurements at five frequencies in the range from 22 to 90 GHz, while the instruments of PLANCK will operate at nine frequencies, covering 20 to 800 GHz. At low frequencies (below 100 GHz) *radio receivers* are used (so called "HEMTs", high electron mobility transistors) while the high frequency instruments are *bolometers*. Recent progress in detector technologies should enable the two new satellites to perform significantly better than COBE – the radio receivers of PLANCK, *e.g.*, are supposed to be 1000 times more sensitive than the ones used for COBE, and the angular resolution has improved from seven degrees to four arc minutes. For more details see

- `http://astro.estec.esa.nl/PLANCK`
- `http://map.gsfc.nasa.gov`
- `http://www.gsfc.nasa.gov/astro/cobe/cobe_home.html`
- `http://spectrum.lbl.gov/www/max.html`
- `http://oberon.roma1.infn.it/boomerang/`

Acknowledgement It is a pleasure to thank Pedro Ferreira, Martin Kunz, Alessandro Melchiorri and Norbert Straumann for stimulationg discussions. I'm also very grateful to Martin Kunz for his help with the preparation of this manuscript.

Appendix A

The C_ℓ's from gravitational waves

We consider metric perturbations which are produced by some isotropic random process (for example during inflation). After production, they evolve according to a deterministic equation of motion. By reasons of isotropy and due to symmetry, the correlation functions of $h_{ij}(\mathbf{k},\eta)$ have to be of the form

$$\begin{aligned}\langle h_{ij}(\mathbf{k},\eta)h^*_{lm}(\mathbf{k},\eta')\rangle &= [k_ik_jk_lk_mH_1(k,\eta,\eta') + \\ &\quad (k_ik_l\delta_{jm} + k_ik_m\delta_{jl} + k_jk_l\delta_{im} + k_jk_m\delta_{il})H_2(k,\eta,\eta') + \\ &\quad k_ik_j\delta_{lm}H_3(k,\eta,\eta') + k_lk_m\delta_{ij}H_3^*(k,\eta',\eta) + \\ &\quad +\delta_{ij}\delta_{lm}H_4(k,\eta,\eta') + (\delta_{il}\delta_{jm} + \delta_{im}\delta_{jl})H_5(k,\eta,\eta')]\ . \end{aligned} \tag{A.1}$$

Here the functions H_a are functions of the modulus $k = |\mathbf{k}|$ only. Furthermore, all of them except H_3 are hermitian in t and t'. This is the most general ansatz for an isotropic correlation tensor satisfying the required symmetries. To project out the tensorial part of this correlation tensor we act on h_{ij} it with the tensor projection operator,

$$T_{ij}{}^{mn} = P_i^mP_j^n - (1/2)P_{ij}P^{mn} \quad \text{with} \quad P_{ij} = \delta_{ij} - \hat{k}_i\hat{k}_j\ . \tag{A.2}$$

This yields

$$\begin{aligned}&\langle h_{ij}^{(T)}(\mathbf{k},\eta)h_{lm}^{(T)*}(\mathbf{k},\eta')\rangle = \\ &\quad H_5(k,\eta,\eta')[\delta_{il}\delta_{jm} + \delta_{im}\delta_{jl} - \delta_{ij}\delta_{lm} + k^{-2}(\delta_{ij}k_lk_m + \\ &\quad \delta_{lm}k_ik_j - \delta_{il}k_jk_m - \delta_{im}k_lk_j - \delta_{jl}k_ik_m - \delta_{jm}k_lk_i) + \\ &\quad k^{-4}k_ik_jk_lk_m]\ . \end{aligned} \tag{A.3}$$

From Eq. (9.17), we then obtain

$$\begin{aligned}&\left\langle \frac{\Delta T}{T}(\mathbf{n})\frac{\Delta T}{T}(\mathbf{n}')\right\rangle \equiv \frac{1}{V}\int d^3x\left(\frac{\Delta T}{T}(\mathbf{n},\mathbf{x})\frac{\Delta T}{T}(\mathbf{n}',\mathbf{x})\right) = \\ &\left(\frac{1}{2\pi}\right)^3\int k^2dkd\Omega_{\hat{\mathbf{k}}}\int_{\eta_{dec}}^{\eta_0}d\eta\int_{\eta_{dec}}^{\eta_0}d\eta'\exp(i\mathbf{k}\cdot\mathbf{n}(\eta_0-\eta))\exp(-i\mathbf{k}\cdot\mathbf{n}(\eta_0-\eta'))\cdot \\ &\left[\langle\dot{h}_{ij}^{(T)}(\eta,\mathbf{k})\dot{h}_{lm}^{(T)*}(\eta',\mathbf{k})\rangle n_in_jn'_ln'_m\right]\ . \end{aligned} \tag{A.4}$$

Here $d\Omega_{\hat{\mathbf{k}}}$ denotes the integral over directions in $\mathbf{k}$ space. We use the normalization of the Fourier transform

$$\hat{f}(\mathbf{k}) = \frac{1}{\sqrt{V}}\int d^3x\exp(i\mathbf{x}\cdot\mathbf{k})f(\mathbf{x})\ , \qquad f(\mathbf{x}) = \frac{1}{(2\pi)^3}\int d^3k\exp(-i\mathbf{x}\cdot\mathbf{k})\hat{f}(\mathbf{k})\ ,$$

where V is an (arbitrary) normalization volume.

We now introduce the form (A.3) of $< h^{(T)}h^{(T)} >$. We further make use of the assumption that the perturbations have been created at some early epoch, e.g. during an inflationary phase, after which they evolved deterministically. The function $H_5(k,\eta,\eta')$ is thus a product of the form

$$H_5(k,\eta,\eta') = H(k,\eta)\cdot H^*(k,\eta')\ . \tag{A.5}$$

Introducing this in Eq. (A.4) yields

$$\begin{aligned}&\left\langle\frac{\Delta T}{T}(\mathbf{n})\frac{\Delta T}{T}(\mathbf{n}')\right\rangle = \\ &\left(\frac{1}{2\pi}\right)^3\int k^2dkd\Omega_{\hat{\mathbf{k}}}\left[(\mathbf{n}\cdot\mathbf{n}')^2-1+\mu'^2+\mu^2-4\mu\mu'(\mathbf{n}\cdot\mathbf{n}')+\mu^2\mu'^2\right]\cdot \\ &\int_{\eta_{dec}}^{\eta_0}d\eta\int_{\eta_{dec}}^{\eta_0}d\eta'\left[\dot{H}(k,\eta)\dot{H}^*(k,\eta')\exp(ik\mu(\eta_0-\eta))\exp(-ik\mu'(\eta_0-\eta'))\right]\ ,\end{aligned} \tag{A.6}$$

where $\mu=(n\cdot\hat{\mathbf{k}})$ and $\mu'=(n'\cdot\hat{\mathbf{k}})$. To proceed, we use the identity [102]

$$\exp((ik\mu(\eta_0-\eta)) = \sum_{r=0}^{\infty}(2r+1)i^r j_r(k(\eta_0-\eta))P_r(\mu)\ . \tag{A.7}$$

Here j_r denotes the spherical Bessel function of order r and P_r is the Legendre polynomial of degree r.

Furthermore, we replace each factor of μ in Eq. (A.6) by a derivative of the exponential $\exp(ik\mu(\eta_0-\eta))$ with respect to $k(\eta_0-\eta)$ and correspondingly with μ'. We then obtain

$$\begin{aligned}&\left\langle\frac{\Delta T}{T}(\mathbf{n})\frac{\Delta T}{T}(\mathbf{n}')\right\rangle = \\ &\left(\frac{1}{2\pi}\right)^3\sum_{r,r'}(2r+1)(2r'+1)i^{(r-r')}\int k^2dkd\Omega_{\hat{\mathbf{k}}}P_r(\mu)P_{r'}(\mu')\times \\ &\Big[2(\mathbf{n}\cdot\mathbf{n}')^2\int d\eta d\eta' j_r(k(\eta_0-\eta))j_{r'}(k(\eta_0-\eta'))\dot{H}(k,\eta)\dot{H}^*(k,\eta') \\ &-\int d\eta d\eta'[j_r(k(\eta_0-\eta))j_{r'}(k(\eta_0-\eta'))+j_r''(k(\eta_0-\eta))j_{r'}(k(\eta_0-\eta'))+ \\ &j_r(k(\eta_0-\eta))j_{r'}''(k(\eta_0-\eta'))-j_r''(k(\eta_0-\eta))j_{r'}''(k(\eta_0-\eta'))]\dot{H}(k,\eta)\dot{H}^*(k,\eta') \\ &-4(\mathbf{n}\cdot\mathbf{n}')\int d\eta d\eta' j_r'(k(\eta_0-\eta))j_{r'}'(k(\eta_0-\eta'))\dot{H}(k,\eta)\dot{H}^*(k,\eta')\Big]\ .\end{aligned} \tag{A.8}$$

Here only the Legendre polynomials, $P_r(\mu)$ and $P_{r'}(\mu')$ depend on the direction $\hat{\mathbf{k}}$. To perform the integration $d\Omega_{\hat{\mathbf{k}}}$, we use the addition theorem for the spherical harmonics Y_{rs},

$$P_r(\mu) = \frac{4\pi}{(2r+1)}\sum_{s=-r}^{r}Y_{rs}(\mathbf{n})Y_{rs}^*(\hat{\mathbf{k}})\ . \tag{A.9}$$

The orthogonality of the spherical harmonics then yields

$$\begin{aligned}&(2r+1)(2r'+1)\int d\Omega_{\hat{\mathbf{k}}}P_r(\mu)P_{r'}(\mu') = \\ &16\pi^2\delta_{rr'}\sum_{s=-r}^{r}Y_{rs}(\mathbf{n})Y_{rs}^*(\mathbf{n}') = \\ &4\pi\delta_{rr'}P_r(\mathbf{n}\cdot\mathbf{n}')\ .\end{aligned} \tag{A.10}$$

In Eq. (A.8) the integration over $d\Omega_{\hat{\mathbf{k}}}$ then leads to terms of the form $(\mathbf{n}\cdot\mathbf{n}')P_r(\mathbf{n}\cdot\mathbf{n}')$ and $(\mathbf{n}\cdot\mathbf{n}')^2P_r(\mathbf{n}\cdot\mathbf{n}')$. To reduce them, we use

$$xP_r(x) = \frac{r+1}{2r+1}P_{r+1}+\frac{r}{2r+1}P_{r-1}\ . \tag{A.11}$$

Applying this and its iteration for $x^2P_r(x)$, we obtain

$$
\begin{aligned}
&\langle\frac{\Delta T}{T}(\mathbf{n})\frac{\Delta T^*}{T}(\mathbf{n}')\rangle = \\
&\frac{1}{2\pi^2}\sum_r(2r+1)\int k^2dk\int d\eta d\eta'\dot{H}(k,\eta)\dot{H}^*(k,\eta')\Big\{ \\
&\left[\frac{2(r+1)(r+2)}{(2r+1)(2r+3)}P_{r+2}+\frac{1}{(2r-1)(2r+3)}P_r+\frac{2r(r-1)}{(2r-1)(2r+1)}P_{r-2}\right]\times \\
&j_r(k(\eta_0-\eta))j_r(k(\eta_0-\eta'))-P_r[j_r(k(\eta_0-\eta))j_r''(k(\eta_0-\eta')) \\
&+j_r(k(\eta_0-\eta'))j_r''(k(\eta_0-\eta))-j_r''(k(\eta_0-\eta))j_{r'}''(k(\eta_0-\eta'))] \\
&-4\left[\frac{r+1}{2r+1}P_{r+1}+\frac{r}{2r+1}P_{r-1}\right]j_r'(k(\eta_0-\eta))j_r'(k(\eta_0-\eta'))\Big\}\ ,
\end{aligned}
\tag{A.12}
$$

where the argument of the Legendre polynomials, $\mathbf{n}\cdot\mathbf{n}'$, has been suppressed. Using the relations

$$
j_r' = -\frac{r+1}{2r+1}j_{r+1}+\frac{r}{2r+1}j_{r-1}
\tag{A.13}
$$

for Bessel functions, and its iteration for j'', we can rewrite Eq. (A.12) in terms of the Bessel functions j_{r-2} to j_{r+2}.

We now insert the definition of C_ℓ:

$$
\left\langle\frac{\Delta T}{T}(\mathbf{n})\cdot\frac{\Delta T}{T}(\mathbf{n}')\right\rangle_{(\mathbf{n}\cdot\mathbf{n}')=\cos\theta} = \frac{1}{4\pi}\Sigma_\ell(2\ell+1)C_\ell P_\ell(\cos\theta)\ ,
\tag{A.14}
$$

and compare the coefficients in Eqs. (A.12) and (A.14). We obtain the somewhat lengthy expression

$$
\begin{aligned}
&C_\ell = \\
&\frac{2}{\pi}\int dkk^2\int d\eta d\eta'\dot{H}(k,\eta)\dot{H}^*(k,\eta')\Big\{j_l(k(\eta_0-\eta))j_l(k(\eta_0-\eta'))\times \\
&\left(\frac{1}{(2\ell-1)(2\ell+3)}+\frac{2(2\ell^2+2\ell-1)}{(2\ell-1)(2\ell+3)}+\frac{(2\ell^2+2\ell-1)^2}{(2\ell-1)^2(2\ell+3)^2}\right. \\
&\left.-\frac{4\ell^3}{(2\ell-1)^2(2\ell+1)}-\frac{4(\ell+1)^3}{(2\ell+1)(2\ell+3)^2}\right) \\
&-[j_\ell(k(\eta_0-\eta))j_{\ell+2}(k(\eta_0-\eta'))+j_{\ell+2}(k(\eta_0-\eta))j_\ell(k(\eta_0-\eta'))]\times \\
&\frac{1}{2l+1}\left(\frac{2(\ell+2)(\ell+1)(2\ell^2+2\ell-1)}{(2\ell-1)(2\ell+3)^2}+\frac{2(\ell+1)(\ell+2)}{(2\ell+3)}-\frac{8(\ell+1)^2(\ell+2)}{(2\ell+3)^2}\right) \\
&-[j_\ell(k(\eta_0-\eta))j_{\ell-2}(k(\eta_0-\eta'))+j_{\ell-2}(k(\eta_0-\eta))j_\ell(k(\eta_0-\eta'))]\times \\
&\frac{1}{2l+1}\left(\frac{2\ell(\ell-1)(2\ell^2+2\ell-1)}{(2\ell-1)^2(2\ell+3)}+\frac{2\ell(\ell-1)}{(2\ell-1)(2}-\frac{8\ell^2(\ell-1)}{(2\ell-1)^2}\right) \\
&+j_{\ell+2}(k(\eta_0-\eta))j_{\ell+2}(k(\eta_0-\eta'))\times \\
&\left(\frac{2(\ell+2)(\ell+1)}{(2\ell+1)(2\ell+3)}-\frac{4(\ell+1)(\ell+2)^2}{(2\ell+1)(2\ell+3)^2}+\frac{(\ell+1)^2(\ell+2)^2}{(2\ell+1)^2(2\ell+3)^2}\right) \\
&+j_{\ell-2}(k(\eta_0-\eta))j_{\ell-2}(k(\eta_0-\eta'))\times \\
&\left(\frac{2\ell(\ell-1)}{(2\ell-1)(2\ell+1)}-\frac{4\ell(\ell-1)^2}{(2\ell-1)^2(2\ell+1)}+\frac{\ell^2(\ell-1)^2}{(2\ell-1)^2(2\ell+1)^2}\right)\Big\}
\end{aligned}
\tag{A.15}
$$

An analysis of the coefficient of each term reveals that the curly bracket in this expression is just

$$
\{\cdots\} = \ell(\ell-1)(\ell+1)(\ell+2)\left(\frac{j_\ell(k(\eta_0-\eta))}{(k(\eta_0-\eta))^2}\right)^2
$$

and the result is equivalent to

$$C_\ell = \frac{2}{\pi} \int dk k^2 |I(\ell, k)|^2 \ell(\ell-1)(\ell+1)(\ell+2) \ , \tag{A.16}$$

with

$$I(\ell, k) = \int_{\eta_{dec}}^{\eta_0} d\eta \dot{H}(\eta, k) \frac{j_\ell((k(\eta_0 - \eta))}{(k(\eta_0 - \eta))^2} \ . \tag{A.17}$$

Bibliography

[1] The Collected Papers of Albert Einstein, Vol. 3, Appendix A, M.J. Klein et al., Eds., Princeton Univ. Press, Princeton, N.J. (1993)

[2] J. Renn, T. Sauer and J. Stachel, The Origin of Gravitational Lensing: A Postscript to Einstein's 1936 Science Paper. Science 275 (1997) 184

[3] F. Zwicky, Phys. Rev. 51 (1937) 290; Phys. Rev. 51 (1937) 679

[4] N. Straumann, General Relativity and Relativistic Astrophysics, Springer, Berlin (1984)

[5] P. Schneider, J. Ehlers and E.E. Falco, Gravitational Lenses, Springer, Berlin (1992)

[6] H. Whitney, Ann. Math. 62 (1955) 374

[7] W. Calley, T. Tyson and E. Turner, Astrophys. J. 461 (1996) L83

[8] M.J. Irvin et al., Astrophys. J. 98 (1989) 1989

[9] J. Binney and S. Tremaine, Galactic Dynamics, Princeton Univ. Press, Princeton (1987)

[10] S. Refsdal and J. Surdej, Rep. Progr. Phys. 56 (1994) 117

[11] R. Narayan and M. Bartelmann, Lectures on Gravitational Lensing, astro-ph 9606001, to appear in the Proceedings of the 13th Jerusalem Winter School of Theoretical Physics (1995)

[12] C.R. Keeton and C.S. Kochanek, in the Proceedings of the IAU Symposium 173 – Astrophysical Applications of Gravitational Lensing, C.S. Kochanek and J.N. Hewitt, Eds., Kluwer, Dordrecht (1996)

[13] B. Fort and Y. Mellier, Astron. Astrophys. Rev. 5 (1994) 239

[14] J.F. Navarro, C.S. Frenk and S.D.M. White, Astrophys. J. 462 (1996) 563

[15] J. Miralda - Escudé and A. Babul, Astrophys. J. 449 (1995) 18

[16] M. Bartelmann and M. Steinmetz, astro-ph 9603101

[17] M. Bartelmann, M. Steinmetz and Weiss, Astron. and Astrophys. 297 (1995) 1

[18] M. Bartelmann, Astron. and Astrophys. 299 (1995) 11

[19] M. Bartelmann, Astron. and Astrophys. 313 (1996) 697

[20] N. Kaiser and G. Squires, Astrophys. J. 404 (1993) 441

[21] N. Kaiser, G. Squires and T. Broadhurst, Astrophys. J. 449 (1995) 460

[22] S. Seitz and P. Schneider, Astron. and Astrophys. 305 (1996) 383

[23] N. Straumann, Helv. Phys. Acta 70 (1997) 894

[24] N. Kaiser, Astrophys. J. 439 (1995) L1

[25] S. Seitz and P. Schneider, MPA report 1077

[26] C. Seitz and P. Schneider, Astron. and Astrophys. 318 (1997) 687

[27] B. Fort and Y. Mellier, Astron. and Astrophys. 5 (1994) 239

[28] C. Seitz, J.P. Kneib, P. Schneider and S. Seitz, Astron. and Astrophys. 314 (1996) 707

[29] N. Kaiser, in the Proceedings of the 18th Texas Symposium on Relativistic Astrophysics, World Scientific, Singapore (1997)

[30] G. Squires et al., Astrophys. J. 482 (1997) 648

[31] S.W. Allen, astro-ph/9710217

[32] M .Hattori, J.-P. Kneib and N. Makino, Progress of Theor. Phys., Supplement no. 133 (1999).

[33] T. Kundić et al., Astrophys. J. 482 (1997) 75

[34] E.A. Carbett, et al., in Ref. [12] 37

[35] L.L.R. Williams and P.L. Schechter, The Quest for the Golden Lens, astro-ph 9709059

[36] M. Fukugita, T. Futamase, M. Kasai and E.L. Turner, Astrophys. J. 393 (1992) 3

[37] M. Im, R.E. Griffiths and K.U. Ratnatunga, Astrophys. J. 475 (1997) 457

[38] M. Chiba and Y. Yoshii, astro-ph/9809321.

[39] C.S. Kochanek et al., Proceedings of the 9th Annual Astrophysics Conference in Maryland, astro-ph/9811111.

[40] Yu-N. Cheng and L.M. Krauss, astro-ph/9810393.

[41] G. Hinshaw and L.M. Krauss, Astrophys. J. **320**, 468 (1987).

[42] S.M. Carroll, W.H. Press and E.L. Turner, Ann. Rev. Astron. Astrophys. **30**, 499 (1992).

[43] M. Bartelmann et al, Astronomy & Astrophys. **300**, 1 (1998).

[44] R. Kaufmann and N. Straumann, astro-ph/9911037.

[45] R. Kaufmann, *Giant Arc Statistics and the Cosmological Parameters*, Diploma Thesis, University of Zürich (1999).

[46] L. Carlson and T.W. Gamelin, Complex Dynamics, Universitext: Tracts in Mathematics, Springer-Verlag, Berlin (1993)

[47] O. Forster, Lectures on Riemann Surfaces, Graduate Texts in Mathematics 11, Springer-Verlag, Berlin (1981)

[48] P. Schneider and C. Seitz, Astron. and Astrophys. 294 (1995) 411

[49] J.B. Conway, Functions of one Complex Variable, Graduate Texts in Mathematics 11, Springer-Verlag, Berlin (1973)

[50] T. Schramm and R. Kayser, Astron. and Astrophys. 299 (1995) 1

[51] T. Schramm, in the Proceedings of the IAU Symposium 173 - Astrophysical Applications of Gravitational Lensing, C.S. Kochanek and J.N. Hewitt, Eds., Kluwer, Dordrecht (1996) 153

[52] J.M. Stewart and M. Walker, Proc. R. Soc. London **A341**, 49 (1974).

[53] J. Bardeen, Phys. Rev. **D22**, 1882 (1980).

[54] H. Kodama and M. Sasaki Prog. Theor. Phys. Suppl **78** (1984).

[55] R. Durrer, Fund. of Cosmic Physics **15**, 209 (1994).

[56] Particle Data Group, Europ. Phys. J. C **3**, 128 (1998).

[57] G.F. Smoot et al., Astrophys. J. **396**, L1 (1992).

[58] R.R. Caldwell, R. Dave and P.J. Steinhardt, Phys. Rev. Lett. **80**, 1582 (1998).

[59] T. Padmanabhan, *Structure formation in the universe*, Cambridge University Press, 1993.

[60] P.J.E. Peebles, *Principles of Physical Cosmology*, Princeton University Press (1993).

[61] E.K. Conklin, Nature **222**, 971 (1969).

[62] P.S. Henry, Nature **231**, 516 (1971).

[63] M. Bruni, P. Dunsby and G. Ellis, Astrophys. J **395**, 34 (1992).

[64] E.M. Lifshitz, Soviet J. Phys. (JETP) **10**, 116 (1946).

[65] P.J.E. Peebles, *Large Scale Structure of the Universe* Princeton University Press (1980).

[66] J.D. Jackson, *Electrodynamics*, Wiley & Sons (New York, 1962, 1975).

[67] S. Chandrasekhar, *Radiative Transfer*, Dover Publications (New York, 1960).

[68] A. Kosowsky, Annals of Physics **246**, 49 (1996).

[69] A. Melchiorri and N. Vittorio, in: Proceedings of the NATO Advanced Study Institute 1996, Strasbourg, archived under `astro-ph/9610029`, (1996).

[70] E.M. Lifshitz and L.P. Pitajewski, *Lehrbuch der Theoretischen Physik, Band X*, Akademie Verlag Berlin (1983).

[71] R.K. Sachs and A.M. Wolfe, Astrophys. J. **147**, 73 (1967).

[72] A.G. Doroshkevich, Ya.B. Zel'dovich, & R.A. Sunyaev, Sov.Astron. **22**, 523 (1978).

[73] E. Harrison, Phys. Rev. **D1** 2726 (1970);
Ya. B. Zel'dovich, Mont. Not. R. Astr. Soc. **160**, P1 (1972).

[74] R. Durrer, A. Gangui and M. Sakellariadou, Phys. Rev. Lett. **76**, 579 (1996).

[75] U. Pen, D. Spergel and N. Turok, Phys Rev. D **49**, 692 (1994).

[76] R. Durrer and Z. Zhou, Phys. Rev. D **53**, 5394 (1996).

[77] B. Allen *et al.*, Phys. Rev. Lett. **79**, 2624 (1997).

[78] R. Durrer *et al.*, Phys. Rev. D **59**, 043511 (1999)

[79] T. Kanazawa, M. Kawasaki, N. Sugiyama, and T. Yanagida, Prog. Theor. Phys. **100** 1055 (1998).

[80] P.J.E. Peebles, Astrophys. J. **510**, 531 (1999).

[81] A. Melchiorri and N. Vittorio, in the proceedings of NATO Advanced Study Institute 1996 on the "Cosmic Background Radiation".

[82] M. Tegmark and A. Hamilton, archived under `astro-ph/9702019` (1997)

[83] K. Ganga, L. Page, E. Cheng and S. Meyer, Astrophys. J. **432**, L15 (1994).

[84] C. M. Gutierrez *et al.*, Astrophys. J. **480**, L83 (1997).

[85] B. Femenia, R. Rebolo, C. Gutierrez, M. Limon and L. Piccirillo, Astrophys. J. **498**, 117 (1998).

[86] K. Coble *et al.*, archived under `astro-ph/9902195` (1999).

[87] G. Tucker *et al.*, Astrophys. J. **475**, 73 (1997).

[88] A. de Oliveira-Costa *et al.* archived under `astro-ph/9808045` (1998).

[89] J.O. Gundersen *et al.*, Astrophys. J. **413**, L1 (1993).

[90] S. Platt *et al.*, Astrophys. J. **475**, L1 (1997).

[91] S. Masi *et al.*, Astrophys. J. **463**, L47 (1996).

[92] B. Netterfield *et al.*, Astrophys. J. **474**, 47 (1997).

[93] Picirillo and P. Calisse, Astrophys. J. **411**, 529 (1993).

[94] A. Miller *et al.*, preprint archived under `astro-ph/9906421`, submitted to Ap.J.Lett. (1999).

[95] E.S. Cheng *et al.*, Astrophys. J. **456**, L71 (1996).

[96] E.S. Cheng *et al.*, Astrophys. J. **488**, L59 (1997).

[97] S.T. Tanaka *et al.*, Astrophys. J. **468**, L81 (1996).

[98] M. Dragovan *et al.*, Astrophys. J. **427**, L67 (1993).

[99] P.F. Scott *et al.*, Astrophys. J. **461**, L1 (1996).

[100] E.M. Leitch *et al.*, preprint archived under `astro-ph/9807312` (1998).

[101] P. de Bernardis *et al.*, Astrophys. J. **422**, L33 (1994).

[102] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, Dover Publications, New York (1972).