M1 Cosmology - The Newtonian Model

Nick Kaiser

November 19, 2021

operational defination
处理得好.

Contents

List of Figures

1 Introduction

1.1 The birth of modern cosmology

Cosmology, as we now understand it, came into existence in a chronologically back-to-front manner. The key events were: 出现

- **Einstein**: relativistic theory of gravity – general relativity (GR) – in 1915.
- **Friedmann**: *expanding relativistic world-model* in 1922
- **Hubble**: discovery (in 1929)
  - that the nebulae he was observing were galaxies exterior to the Milky Way
  - and that they were receding from us with velocity proportional to distance

1.2 The FLRW models

Lemaitre worked on applying GR to cosmology a little later than Friedmann. And in the 30's **Robertson** and **Walker** elucidated many features of these models, which nowadays go by the name of **FLRW models**.

It was established how observations of galaxy flux-densities (apparent luminosities) and redshifting of spectral lines in galaxy spectra can be used to constrain the parameters of the model, which can be taken to be the current *expansion rate* and the rate at which this is changing – the *deceleration parameter*.[1]

These models, dating back now nearly 100 years, are those which are still used by practising cosmologists to interpret observations.

1.3 Physical interpretation of the FLRW models

At the time of Friedmann, and even up to the 60s, the understanding of GR was in a process of development. The *Schwarzschild solution* for a black-hole, for example, was found in 1916, but the physical meaning of the *gravitational radius* $r = 2GM/c^2$ was not fully appreciated for decades. In cosmology, the interpretation of the FLRW models was also problematic, particularly in relation to the reason for the cosmological redshift. One finds, in many text-books and articles, statements about the *'expansion of space'* and how this causes light to be redshifted. One reads, in Harrison (2000) for instance, that *"expansion redshifts are produced by expansion of space between bodies that are stationary in space"*. In many other works, the fact that light is redshifted in an expanding universe is held to be self-evident. The FLRW metric contains the expansion factor $a(\tau)$, and the metric plays the role, in GR, of the potential in Newtonian gravity. If one formulates Maxwell's equations in FLRW coordinates, one finds a term (containing the Hubble expansion rate $H = \dot{a}/a$) that is widely interpreted as expressing the *'coupling of electromagnetism to the gravitational field'*. This is not entirely crazy. In weak-field gravity, for instance, the metric is $g_{\alpha\beta} = \eta_{\alpha\beta} - 2\phi\delta_{\alpha\beta}/c^2$ where ϕ is the Newtonian potential. In FLRW models, the metric contains the scale factor, so it might not seem unreasonable to describe effect of the cosmological expansion as a coupling to gravity.

But, on closer inspection, this falls apart. The gravitational field in GR is the *tide* – or the curvature of space-time – not the expansion. In the foundations of GR, there is nothing one can really identify as the 'expansion of space' *per se*; all there are are measurable distances, from which one can distill the metric (and which relates these to coordinate separations – which are arbitrary). Distances we measure between galaxies are increasing, and the amount of space in a volume enclosed by a set of galaxies is increasing, and indeed the total amount of space in a closed FRW model is a well defined and is, in general, changing. But nowhere in GR do we find any way of actually measuring the expansion of space *itself*. Analogies are often drawn with expanding rubber balloons, but the expansion of a balloon is something one can measure; mark some points and then measure their distance with a ruler. In GR there is no way to anchor objects to space-time. Indeed, the principle on which the theory is based is that locally space-time is *Minkowskian*. Such space-time has an absolute sense of *rotation* – if you are rotating with respect to it then you can feel it – but it has no sense of expansion. In cosmology 'stuff' – including radiation – is expanding, but space, of itself, is not. The idea that the expansion of space – or, as it is often said, the 'fabric of space-time' – is a real

布料；构造

[1] Alan Sandage, who worked, as a student, with Hubble, and continued Hubble's work after his death in 1953, famously characterised cosmology as *"the search for two numbers"*.

险恶的

phenomenon with local physically measurable effects, is a pernicious one. In 1945, Einstein wrote a paper with Ernst Straus entitled *'The Influence of the Expansion of Space on the Gravitation Fields Surrounding the Individual Stars'*. They concluded that there is none, but that did not dispel the myth.

So there is a lot of historical baggage and misconceptions in the cosmology literature. One motivation for considering Newtonian cosmology – which works perhaps surprisingly well – is that it helps to avoid falling into some of these traps. In regard to the role of the expansion in affecting fields and particle motions, it shows that the appearance of the additional 'Hubble damping term' is better considered not to be a physical effect but simply a 'coordinate artefact'.

1.4 The Newtonian analogy

In 1934 **E.A. Milne**, in two papers, one with **Bill McCrea**, noted the close resemblance of the Friedmann equations to those of Newtonian dynamics and stated that *"All of the phenomena observable at the present could have been predicted by the founders of mathematical hydrodynamics in the 18th century, or even by* **Newton** *himself"*.

In the rather confused context of their time, Milne and McCrea's Newtonian description of cosmology is a useful one. It works because, in the FLRW models the high degree of symmetry means the dynamics is purely local. While, in general, one can measure the tidal field due to distant matter – as we see the tidal influence of the moon on the oceans, for example – in the FLRW models there is none. The only tidal field is the local isotropic focusing of particles due to the attraction of the local matter – or repulsion if we are dealing with quintessence or the cosmological constant. And the local dynamics involves velocities that are small, so relativistic effects associated with the expansion velocities are negligible. And, as we shall see, one can include the effect of relativistic particles (i.e. the radiation) by means of special relativity.

But it does not completely replace GR. What is missing is curvature. When Milne and McCrea wrote, 'the phenomena observable at the present' meant observations of objects at relatively low redshift, and curvature only shows up when one goes beyond linear order in z (in the apparent luminosity and angular diameter distances $D_{\rm a}$ and D_L say).

It is also useful in the context of *modified theories of gravity* to have a clear idea of what aspects of cosmology would not change under such modifications (assuming that, like GR, they reduce to Newtonian gravity in the appropriate limit).

1.5 Why didn't Newton do Newtonian cosmology?

Before leaping into the mathematics of Newtonian cosmology one may reasonably ask the question above. While not much cosmology appears in the *Principia* or in *Opticks*, there is correspondence between Newton and the English philosopher **Richard Bentley**, in which he discusses the behaviour of an infinite distribution of gravitating masses.

While he made progress in understanding what we would call today *gravitational instability* – the process by which mass would aggregate into structures – he was less successful in developing what we would call the *background model*[2].

Bentley suggested that an infinite sea of stars – if initially at rest – would collapse to a central point. Since there is no 'centre' for an infinite field of stars, this seems paradoxical. Newton resisted this and 'solved' the paradox by claiming that God prevented the collapse by making "constant minute corrections", though to be fair to him, he was not entirely happy with this divine intervention.

What Milne and McCrea pointed out is that had Newton considered the dynamics of an initially expanding sphere of 'dust' (the dust being stars or galaxies; i.e. pressure-less matter) of a finite radius, he would have seen that the radius plays no essential role in the dynamical equations. Thus, in Newtonian theory, in which velocities can be arbitrarily large, the radius can be arbitrarily large, he would have obtained a satisfactory solution for an infinite expanding universe, some two centuries ahead of Friedmann.

[2] In working on the former, in the absence of the latter, he was applying what nowadays goes by the name of *Jeans' swindle*

2 Radial orbits in the field of a point mass

2.1 The equation of motion

A particle on a radial orbit outside a mass M has acceleration

$$\boxed{\ddot{r} = -GM/r^2} \tag{1}$$

where G is Newton's gravitational constant, and dot denotes time derivative.

A consequence of this is that (twice) the energy per unit particle mass is

$$K \equiv \dot{r}^2 - 2GM/r = \text{constant.} \tag{2}$$

To confirm this, note that $dK/dt = d/dt(\dot{r}^2 - 2GM/r) = 2\dot{r}\ddot{r} + 2GM\dot{r}/r^2 = 2\dot{r} \times (\ddot{r} + GM/r^2) = 0$.

2.2 Parametric (cycloid and hyper-cycloid) solution

Despite the simple form of the acceleration equation (1), there is no analytic formula for its general solution. There is, however, a parametric solution. For the case of a bound particle, is the cycloid

$$\begin{aligned} r(\eta) &= A(1 - \cos\eta) \\ t(\eta) &= B(\eta - \sin\eta) \end{aligned} \tag{3}$$

where A and B are constants.

The reason it is called a cycloid is that it is equation for the path of a particle on the rim of a rolling bicycle wheel. For a wheel of unit radius, the path of the axle is horizontal and given by: $(x_0, y_0) = (\eta, 1)$, with η being the angle of rotation about the axle.
A particle on the rim has trajectory given by $(x, y) = (x_0, y_0) + (\sin\eta, \cos\eta)$, the second term being the displacement of a point on the rim from the axle, so

$$\begin{aligned} y &= 1 - \cos\eta \\ x &= \eta - \sin\eta \end{aligned} \tag{4}$$

It is straightforward to prove that (3) solves $\ddot{r} = -GM/r^2$. The derivatives with respect to conformal time η (denoted by primes) are $r' = dr/d\eta = A\sin\eta$ and $t' = dt/d\eta = B(1 - \cos\eta)$ so $\dot{r}^2 = (r'/t')^2 = (A/B)^2 \sin^2\eta/(1-\cos\eta)^2$, from which we can eliminate $\sin\eta$ in favour of $1-\cos\eta$ using $\sin^2\eta = 1 - \cos^2\eta = (1-\cos\eta)(1+\cos\eta) = (1-\cos\eta)(2-(1-\cos\eta))$ so

$$\begin{aligned} \dot{r}^2 &= \frac{A^2}{B^2}[2/(1-\cos\eta) - 1] \\ &= \frac{2A^3}{B^2}\frac{1}{r} - \frac{A^2}{B^2} \end{aligned} \tag{5}$$

which agrees with $\dot{r}^2 = 2GM/r + K$ if $A^2 = -KB^2$ and $AK = -GM$ or

$$\begin{aligned} A &= -GM/K \\ B &= GM/|K|^{3/2} \end{aligned} \tag{6}$$

This solution only applies if total energy K is negative – i.e. bound orbit. For such an orbit the particle reaches maximum radius $r_{\rm max} = 2GM/|K|$ at time $t_{\rm max} = t(\eta = \pi) = \pi GM/|K|^{3/2}$ and falls back to $r = 0$ at $t = t(\eta = 2\pi) = 2t_{\rm max}$.

On the other hand, if the particle has $K > 0$ – i.e. velocity exceeding the *escape velocity* – solution is a hyper-cycloid, in which we simply replace sin $\rightarrow$ sinh and, paying regard to signs, we get

$$\begin{aligned} r(\eta) &= \frac{GM}{K}(\cosh\eta - 1) \\ t(\eta) &= \frac{GM}{K^{3/2}}(\sinh\eta - \eta) \end{aligned} \tag{7}$$

At late time ($\eta \gg 1$): $r \simeq (GM/K)e^{\eta}/2$ and $t \simeq (GM/K^{3/2})e^{\eta}/2$, so asymptotically $r = t/\sqrt{K}$.

The cycloid and hyper-cycloid solutions are (for fixed central mass M) a family parameterised by (twice) the total energy per unit test-particle mass K. They can also be thought of as being parameterised by k (equal to +1 for the cycloid and -1 for the hyper-cycloids) and the time $B = GM/|K|^{3/2}$. They are shown in figure 1.

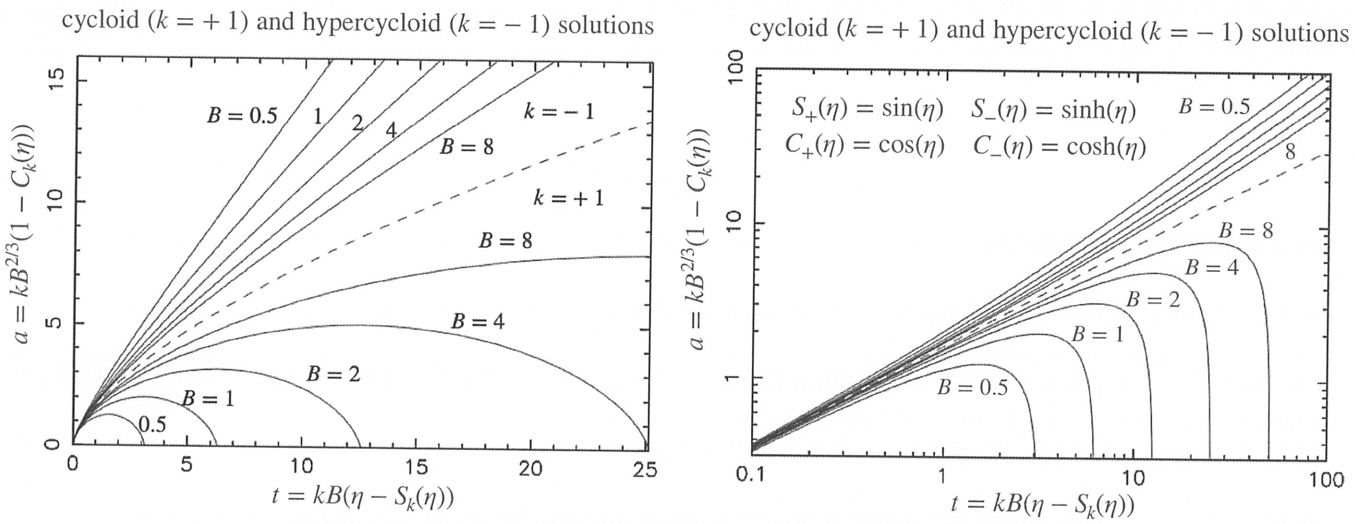


Figure 1: Solutions for the motion of a particle launched on a radial orbit in the gravitational field of point mass. The variable $B = GM/|K|^{3/2}$ has units of time and for the bound orbits ($k = +1$) is equal to $(r_{\max}^3/8GM)^{1/2}$. The ordinate here is $a = r/(GM)^{1/3}$. Left plot is linear-linear. Right plot is the same thing but on a log-log scale. The dashed line is the marginally bound Einstein-de Sitter solution, which is the limit as $B \to \infty$ of either the cycloid or hyper-cycloid. At early times the curves become indistinguishable as the kinetic and potential energy terms both become very large compared to their difference and the orbits are, asymptotically $r(t) = (9GM/2)^{1/3}t^{2/3}$.

2.3 Behaviour at early times

At early times (i.e. $\eta \ll 1$), and for both cycoid and hyper-cycloid, the energy constant becomes negligible (compared to $\dot{r}^2$ or $2GM/r$)

Taylor expanding the solutions for $\eta \ll 1$ we have

$$\begin{aligned} r &= (GM/K)(\eta^2/2 + \mathcal{O}(\eta^4)) \\ t &= (GM/K^{3/2})(\eta^3/6 + \mathcal{O}(\eta^5)) \end{aligned} \tag{8}$$

so

$$\boxed{r = \left(\frac{9GM}{2}\right)^{1/3} t^{2/3}} \tag{9}$$

for $\eta \ll 1$ (or, equivalently, for $t \ll GM/|K|^{3/2}$) and, defining density by $\rho = 3M/4\pi r^3$, time is related to density by

$$\boxed{G\rho = \frac{1}{6\pi t^2}} \tag{10}$$

in accord with the usual relationship between dynamical (or orbital) time and density $t_{\rm dyn} \sim 1/\sqrt{G\rho}$. These relations apply at all times in the marginally bound case

3 A uniform density expanding dust sphere

Consider an initially uniform density sphere of 'dust' that is expanding with velocity proportional to distance from the centre: $\dot{r} = Hr$. A particle at the edge has equation of motion – the *acceleration equation*:

$$\ddot{r} = -GM/r^2 \tag{11}$$

and so has a (hyper-)cycloidal orbit.

But, according to Newton, a particle at some smaller radius only 'feels' the acceleration from the mass interior and that mass scales, initially, as r^3, so $GM/r^2 = (4\pi/3)G\rho r$, which is linear in r, so, after an interval of times dt the velocity will change by $d\dot{r} = \ddot{r}dt = -(4\pi/3)G\rho r dt$ and hence the fractional change in the velocity is $d\dot{r}/\dot{r} = -(4\pi G\rho/3H)dt$ which in independent of r.

So, in this interval, all the shells change their radius by the same fractional amount. Consequently the shells remain nested in exactly the same order and the density remains uniform. What's more, the velocity still retains its pure Hubble-law form.

It follows that all of the shells have equation of motion

$$\boxed{\ddot{r} = -(4\pi/3)G\rho r} \tag{12}$$

independent of their radius.

- Note that it was critical here to assume an inverse square attraction law.

So the time dependence of the solution is the same for all shells independent of r, and this implies that *an initially uniform density sphere remains uniform.*

What's more, *observers inside the sphere can't tell where the centre is.*

- Q: doesn't gravity vector $\mathbf{g} = -GM\mathbf{r}/r^3 = -(4\pi/3)G\rho\mathbf{r}$ break the symmetry?
- A: for an electrically charged sphere it would since we can measure $\mathbf{E}$ (the analogue of $\mathbf{g}$) by observing motion of two test particles with different charge-to-mass ratio. But according to Galileo, in gravity, all particles have same 'charge-to-mass' so there is no way, using motions of particles, to measure $\mathbf{g}$.
- only the 'tidal field' – i.e. the way $\mathbf{g}$ changes with position – is measurable.

that means we can make the dust sphere as large as we like, and in Newtonian physics velocities can be arbitrarily large, and all points within the sphere are equivalent.

All there is is a local expansion rate and local density which are the same at all points in space – and the 'tide' is purely radial; any pair of test particles accelerate towards themselves in proportion to their separation and the local density.

The 'Galilean equivalence principle' has effectively rendered the 'boundary conditions at infinity' – the very distant edge of the sphere – unobservable.

- note, however, that this is only for a sphere; had we considered a uniformly expanding *cube*, for instance, we would have observable effects from the boundary. The cube would not remain cubical, for instance, and the expansion rate would not be isotropic.

To summarise:

- within the (arbitrarily large) sphere all points are equivalent with regard to local dynamics
- this being a special property of the inverse square attraction law
- and the equivalence principle played a direct and central role in making it possible to construct a well-behaved cosmological model
  - if one were tempted to try to construct a homogeneous and isotropic model in which the expansion of the universe were a result of it having a net electrical charge, one would immediately face a severe problem

This is all something that Newton could have figured out. He certainly thought about it, and discussed it, but he apparently he got it wrong. He realised what would happen for a finite sphere, but somehow convinced himself that an infinite sphere would not behave in the same manner.

4 Friedmann, continuity and acceleration equations

4.1 Re-scaled or 'comoving' coordinates

Thanks to the spatial uniformity can re-scale the spatial coordinates for the dust particles according to

$$\mathbf{r} = a(t)\mathbf{x} \tag{13}$$

where

- $\mathbf{x}$ is a dimensionless *'co-moving' coordinate*
  - effectively a label – that is fixed for each particle
- and $a(t)$ is a *universal scale factor*
  - it is the distance between a pair of particles that have unit comoving separation $|\mathbf{x} - \mathbf{x}'| = 1$

4.2 The re-scaled energy equation

The energy equation is $\dot{r}^2 = (8\pi/3)G\rho r^2 - K$, where the constant K is, in general, different for different particles.

But since $\dot{x} = 0$, $\dot{r} = x\dot{a}$ so both the first two terms scale in proportion to x^2. Thus K must must be proportional to x^2 also.

Now K has units of velocity squared, so we can write $K = kc^2x^2$ where k is a dimensionless constant.

Taking out the common factor x^2 gives the *re-scaled energy equation*

$$\boxed{\dot{a}^2 = (8\pi/3)G\rho a^2 - kc^2} \tag{14}$$

which, as we will see later, is identical *in form* to the general relativistic *Friedmann equation*.

But there k was the 'curvature constant' and was limited to be $k = -1, 0, 1$, whereas here k is arbitrary.

But so is the labelling of the particles, and hence the value of a (at some chosen time) so we could, if we like, and will require $k = -1, 0, 1$.

In the relativistic framework, a is the separation between two particles that, asymptotically at late times (in the hyperbolic case), increases at the speed of light.

Dividing by a^2 and gives the *Friedmann (energy) equation* in the form

$$\boxed{(\dot{a}/a)^2 = (8\pi/3)G\rho - kc^2/a^2} \tag{15}$$

just as in the relativistic FRW models.

While conservation of mass implies that the density varies as $\rho \propto 1/a^3$ which implies the *continuity equation*

$$\boxed{\dot{\rho} = -3(\dot{a}/a)\rho} \tag{16}$$

while differentiating the Friedmann equation and using continuity gives the *acceleration equation*

$$\boxed{\ddot{a}/a = -(4\pi/3)G\rho} \tag{17}$$

which is what we started with.

These three equations are not independent – any pair of them implies the third, and any pair provides two equations for the two functions of time $a(t)$ and $\rho(t)$.

As we will see later, the Newtonian formulae above are identical to the relativistic equations (for $P = 0$).

The reason, ultimately, that the Newtonian and relativistic treatments agree is that the form of the latter is fixed by the fact that they need to properly describe the local expansion, which, thanks to Gauss's law, is independent of what is happening outside the local region

4.3 The expansion rate

The *expansion rate* is defined by

$$\boxed{H \equiv \frac{\dot{a}}{a}} \tag{18}$$

and has units of inverse time. Its value at the present epoch is usually denoted by H_0 and is called the *Hubble parameter* or *Hubble's constant*. In the units that cosmologists like to use, $H_0 \simeq 70\text{km/sec/Mpc}$.

5 Cosmological observables in the Newtonian model

5.1 Redshift: Peebles's argument

One of the key observables in cosmology is the redshift of light from distant objects. This is most simply understood as being the accumulation of a series of small Doppler shifts as would be perceived by a set of observers riding with galaxies along the light path. Newton would probably have had some trouble with this since he believed in the 'corpuscular' theory of light[3]. (微粒的)

The finiteness of the speed of light – measured to $\sim 20\%$ precision by Ole Rømer in 1676 – was however already known to Newton by the time he came up with his universal law of gravitation (published in 1686). So he would have been able to calculate the manner in which the observed intervals between emitted light pulses would be dilated by the expansion of the cloud of observers, which amounts to the same thing as the dilation of wave-length (in a wave-theory for light).

One might object this would require special relativity to give the Doppler shift. But in fact, as we consider a sequence of small intervals, it is sufficient to use the 1st order Doppler shift formula $\delta\lambda/\lambda = \delta v/c$. This is non-relativistic, in that we ignore the second order terms and the transverse Doppler shift, but does depend on the finite speed of light.

Armed only with this, the redshift – change of wavelength at measured by an expanding cloud of observers (so-called *fundamental observers*) – of light can be shown to be the same as the ratio of the scale factor at reception to that as emission by considering it to be the product of an infinite number of infinitesimal shifts, as explained by Peebles.

Considering the infinitesimal Doppler shift suffered by light travelling between two particles with infinitesimal comoving separation dx

- They have physical separation $dr = adx$
- and relative recession velocity $dv = \dot{a}dx$
- so the fractional change in wavelength is $d\lambda/\lambda = dv/c = \dot{a}dx/c$
- but $\dot{a} \equiv da/dt$ and the time interval is $dt = adx/c$
- so $d\lambda/\lambda = da/a$
- which implies $\lambda \propto a$

thus

$$\boxed{1 + z \equiv \frac{\lambda_{\rm obs}}{\lambda_{\rm em}} = \frac{a(t_{\rm obs})}{a(t_{\rm em})}} \tag{19}$$

A useful alternative way of thinking about this is in terms of differences in time at the emitter and observer between pulses or light signals:

$$\boxed{1 + z = \frac{dt_{\rm obs}}{dt_{\rm em}}.} \tag{20}$$

5.2 The redshift as a combination of a Doppler and gravitational effects

You might worry that there might be, in addition to the 1st order Doppler shift, a *gravitational redshift*, but you don't need to. Think of the light propagating from an observer at $\mathbf{x} = 0$ to another observer with an infinitesimal separation $|\mathbf{x}| = dx$. The gravitational redshift is the difference in the potential, but the potential, for a uniform density, is quadratic, so the gravitational redshift is $\sim G\rho dx^2$. When you consider the finite path wavelength ratio as the limit as $dx \to 0$ of the product of a lot of infinitesimal steps this does not contribute; only the 1st order Doppler effects survive.

Similarly, for an infinitesimal separation you don't need to worry about the fact that we were a little loose in saying the relative recession velocity is $dv = \dot{a}dx$. What is $\dot{a}$ here? the value at the beginning of the trip? At the end? Some kind of average? It doesn't matter as it only affects the answer for an infinitesimal interval at 2nd order.

[3]Query 29 of *'Opticks'* is *"Are not the Rays of Light very small Bodies emitted from shining Substances?"*

terminology: brightness = surface brightness
flux density,

Einstein (1910) thought experiment

The gravitational redshift experiment. **Let us first imagine performing an idealized experiment, first suggested by Einstein. (i) Let a tower of height h be constructed on the surface of Earth, as in Fig. 5.1. Begin with a**

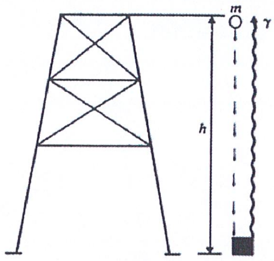
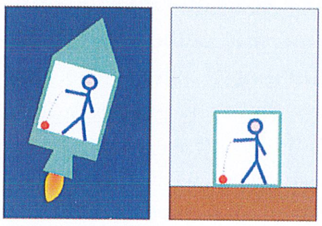
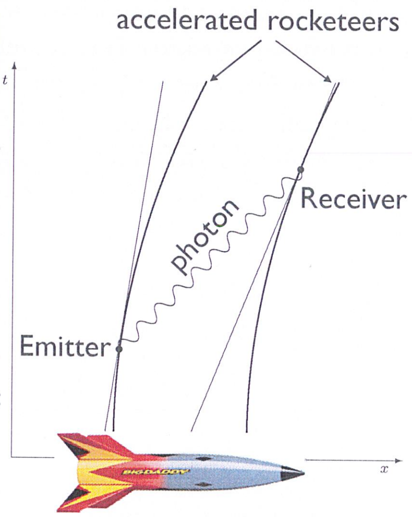


Fig. 5.1 A mass m is dropped from a tower of height h. The total mass at the bottom is converted into energy and returned to the top as a photon. Perpetual motion will be performed unless the photon loses as much energy in climbing as the mass gained in falling. Light is therefore redshifted as it climbs in a gravitational field.

Figure 2: In Einstein's tower thought experiment he imagines a special machine – not one that actually exists, but one that obeys the laws of physics – that receives a mass dropped from to top of the tower and converts its energy E – rest-mass plus kinetic – to a photon of frequency $\nu = E/h$ that is fired back to an identical machine at the top that converts the photon back to a stationary mass (again conserving energy). Without the gravitational redshift this would be a way to continuously extract useful energy from a static gravitational field.

Einstein's calculation of the redshift in a rocket





- during time $\delta t = x / c$ it takes the photon to make trip the velocity of receiver changes: $\delta v = g\, \delta t = g\, x / c$.
- Doppler shift: $\delta\lambda/\lambda = \delta v / c = g\, x / c^2$
- But is this gravity?

Figure 3: Einstein's rocket thought experiment. He considers a photon fired from an observer at the back of an accelerating rocket to another at the front. He argues that there would be a 1st order Doppler shift what would be the same as the frequency shift for a pair of non-accelerating observers (straight lines) who are co-moving with the accelerating observers at the times of emission and reception.

However, if we consider an expanding dust sphere of *finite* size then, as first realised by Bondi, one can think of the redshift as being a combination of Doppler and gravitational shifts (at least for modest redshifts).

But as shown in figure 3, the combination ends up being solely the change in the separation between the source and observer and the gravitational redshift is effectively hidden

5.3 Comoving-distance vs. redshift relation

What is the relation between the redshift of a source z and x, its comoving distance from us?

- in a interval dt light travels $dr = -cdt$ so $dx = dr/a = -cdt/a$
- while $1 + z = a_0/a$ (where subscript "0" denotes present value)
- so $dz = -(a_0/a^2)da = -a_0\dot{a}dt/a^2 = -(a_0/a)Hdt = (a_0 H/c)dx$ and hence

$$dx = \frac{c}{a_0}\frac{dz}{H(z)} \tag{21}$$

with integral

$$x(z) = \frac{c}{a_0}\int_0^z \frac{dz}{H(z)} \tag{22}$$

which, as we will see later, is the same is the relativistic formula for something called the *conformal distance* χ.

To evaluate $x(z)$:

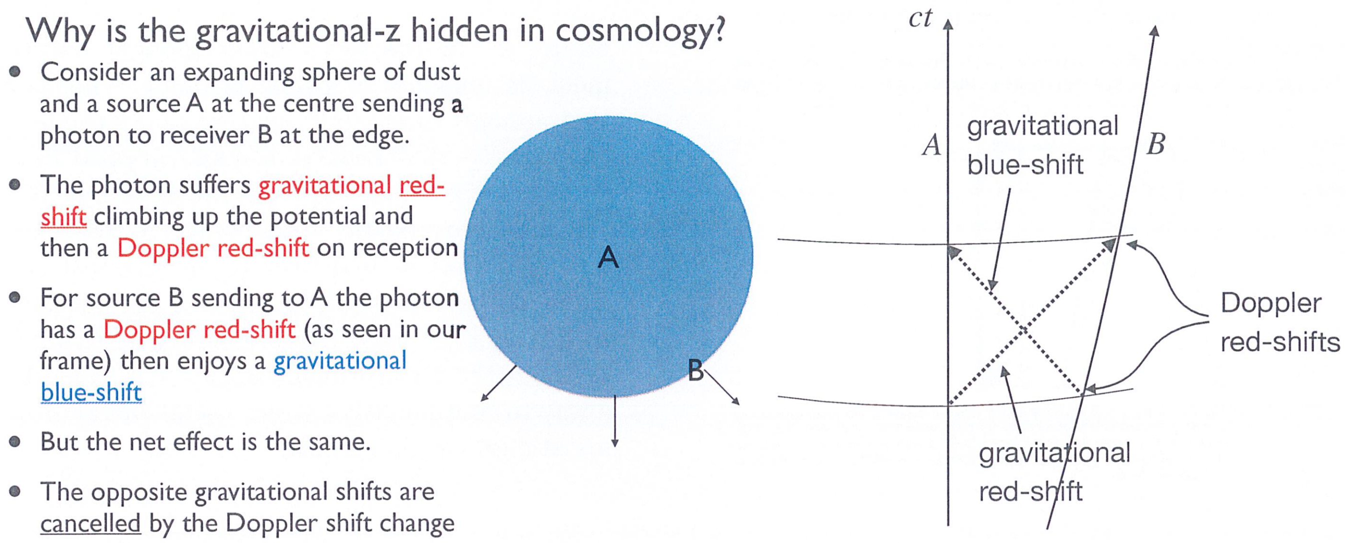


Figure 4: Bondi showed that, for sources at low redshift at least, one can think of the cosmological redshift as being a combination of a Doppler shift and a gravitational redshift. But somehow this comes out to be just the change in the separation between the source and observer – a purely 'kinematic' effect. So, for instance, in the situation shown at left, where if a photon is sent from A to B it suffers a gravitational redshift while for the opposite direction it enjoys a blue-shift, the difference is cancelled by a corresponding change in the Doppler component. This is a consequence of the highly symmetric situation.

- define *critical density* $\rho_c \equiv 3H_0^2/8\pi G$
- and *density parameter* $\Omega_m \equiv \rho_0/\rho_c$
- and let $\Omega_k \equiv 1 - \Omega_m$

then the Friedmann energy equation is

$$H^2(z) = H_0^2[\Omega_m(1+z)^3 + \Omega_k(1+z)^2] \tag{23}$$

so the comoving distance of a source observed to have redshift z is

$$x(z) = \frac{c}{H_0 a_0} \int_0^z dz / \sqrt{\Omega_m(1+z)^3 + \Omega_k(1+z)^2} \tag{24}$$

so the *physical* distance – *at the present epoch* – is

$$a_0 x(z) = \frac{c}{H_0} \int_0^z dz / \sqrt{\Omega_m(1+z)^3 + \Omega_k(1+z)^2} \tag{25}$$

which, as we will see later, is the same as the relativistic result (with $x \rightarrow \chi$)

An interesting feature of the comoving distance in this model is that $x(z)$ remains finite as source redshift $z \rightarrow \infty$. This means that there is a *horizon*; all the sources we can see – which must have finite redshift – lie in a finite volume of space.

5.4 Angular diameter and luminosity distances in Newtonian cosmology

The comoving distance x is not something that observers can easily measure. What they can measure is the angular size θ of distant objects and their apparent luminosities (or flux densities F).

If one has objects of known, or standard, size d, it is useful to define the *angular diameter distance* $D_a(z)$ such that $\theta = d/D_a(z)$.

And if one has objects of known, or standard, luminosity L, it is useful to define the *luminosity distance* $D_L(z)$ such that the flux density is $F = L/4\pi D_L^2$.

名义距离

These are *apparent distances*. The angular diameter distance is the answer to the question: how far away from us would the object (of assumed known intrinsic size d) need to be to have the angular size we observe. The luminosity distance is the distance a source of (assumed known intrinsic luminosity L) would have to be at – in a non-expanding universe – to have the flux-density we observe.

The reason that these quantities are of interest, is that one can calculate them using the formulae above, and the results depend on the parameters H_0 and $\Omega_{\rm m}$. So if one had some way of figuring out the intrinsic size or luminosity (the latter being possible with type 1a with supernovae), one can determine these parameters.

If Newton had succeeded in developing Newtonian cosmology, what would he have had to say about the apparent distances $D_{\rm a}(z)$ and $D_L(z)$?

The answer to this is somewhat tricky. With regard to $D_{\rm a}$, he would certainly have pondered on the influence of light deflection. In the 3rd of his trilogy *'Opticks'* he has a series of 'queries', the first of which is *"Do not Bodies act upon Light at a distance, and by their action bend its Rays, and is not this action (cæteris paribus) strongest at the least distance?"*.

So he is here predicting the deflection of light by masses – and getting it wrong by only a factor two – but what would he have said about light paths in a large – potentially very large - expanding uniform density sphere?

It seems to me that, just as he struggled with the transition from a large finite sphere to an infinite one (concluding that in the latter case the matter would remain at rest), he would have argued that, for a spherical distribution of matter, the light rays would, by symmetry, have to be undeflected.

Thus, he would presumably have thought that the angular size of a 'standard ruler' is it's physical size d divided by its physical distance *at the time that the photons left the source*. That would lead to $D_{\rm a} = a(t_{\rm em})x$. Replacing $a(t_{\rm em}$ by $a_0/(1+z)$ we have

$$D_{\rm a}(z) = \frac{a_0 x}{1+z} = \frac{c}{1+z}\int_0^z \frac{dz}{H(z)}. \tag{26}$$

Remarkably, this is in agreement with the relativistic formula in the case that the universe is spatially flat. As you will see in the travaux dirigés, the angular diameter distance increases with redshift at low redshift, but reaches a maximum and then decreases.

As regards the luminosity distance he would similarly have argued that for source emitting $dn/dt_{\rm em} = N$ photons (light corpuscles) per unit time at the source the number of photons per unit time crossing a sphere (on which reside observers who see the source to have redshift z) would be $N \times t_{\rm em}/t_{\rm obs} = N/(1+z)$. The area of that sphere is $A = 4\pi(a_0 x(z))^2$ so the flux density of photons at the observer is $N/A = N/[4\pi(a_0 x(z))^2(1+z)]$. Equating that to the photon number flux density $N/4\pi D_N^2$ one would see for the same source at fixed distance D in empty space gives

$$D_N = a_0\sqrt{1+z}x = (1+z)^{3/2} D_{\rm a}. \tag{27}$$

Finally, using special relativity to argue that the energy of the photons is reduded by the redshift would give energy flux-density $F = L/[4\pi(a_0 x(z))^2(1+z)^2]$ so, on equating $F = L/4\pi D_L^2$ gives

$$D_L = a_0(1+z)x = (1+z)^2 D_{\rm a}. \tag{28}$$

The relation above between D_L and D_A is called *Etherington's reciprocity relation.*

One should beware that the luminosity distance obtained above is strictly valid for so-called *bolometric* flux densities where all the photons are detected and the total energy measured. In practice, which is more often measured are flux densities measured with a *broad-band filter* in front of the detector, and, in the case of CCD detectors which are most commonly used, the measured quantity is a number of photo-electrons. In general, the appropriate apparent distance must be calculated by performing an integral over wavelength, and, again in general, one needs to know the spectrum – or *spectral energy density* SED – to calculate this. A good approximation may be obtained by measuring flux densities in a number of pass-bands, and using the 'colour' obtained from the different pass-bands to make a correction to the luminosity distance. This is called *applying a K-correction.*

These apparent distances one would obtain in the Newtonian picture – assuming light rays are undeflected – are equivalent to the correct relativistic expressions, but only in the case that the universe is spatially flat ($k = 0$).

This is a real restriction on the Newtonian model. However, perhaps luckily, it turns out that our universe is, or is very close to being, spatially flat.

6 Structure formation in Newtonian cosmology

6.1 The Dmitriev & Zel'dovich equations

For N particles of mass m interacting under their mutual gravitational attraction Newton's laws of motion are $3N$ second order differential equations

$$\ddot{\mathbf{r}}_i = Gm \sum_{j \neq i} \frac{\mathbf{r}_j - \mathbf{r}_i}{|\mathbf{r}_j - \mathbf{r}_i|^3}. \tag{29}$$

These may be solved numerically provided initial positions $\mathbf{r}_i$ and velocities $\dot{\mathbf{r}}_i$ for the particles.

Writing this in terms of arbitrarily re-scaled coordinates $\mathbf{r} = a(t)\mathbf{x}$, so $\dot{\mathbf{r}} = \dot{a}\mathbf{x} + a\dot{\mathbf{x}}$ and $\ddot{\mathbf{r}} = \ddot{a}\mathbf{x} + 2\dot{a}\dot{\mathbf{x}} + a\ddot{\mathbf{x}}$, (29) becomes

$$\ddot{\mathbf{x}}_i + 2\frac{\dot{a}}{a}\dot{\mathbf{x}}_i = \frac{Gm}{a^3} \sum_{j \neq i} \frac{\mathbf{x}_j - \mathbf{x}_i}{|\mathbf{x}_j - \mathbf{x}_i|^3} - \frac{\ddot{a}}{a}\mathbf{x}_i \tag{30}$$

where we have, somewhat arbitrarily, moved one of the terms in $\ddot{\mathbf{r}}$ over to the right hand side.

What we are interested in is the motion of particles with initial conditions that are close to being in uniform Hubble expansion with some initial expansion rate H (very close if we start at early times). So we might lay down particles on a regular grid in $\mathbf{r}$-space within some very large spherical boundary centred on the origin and give the particles small displacements $\delta\mathbf{r}$ and velocities $\dot{\mathbf{r}} = H\mathbf{r} + \delta\dot{\mathbf{r}}$ with 'peculiar' velocities $\delta\dot{\mathbf{r}}$ chosen to excite the growing mode (see below). This is illustrated in figure 4. The corresponding initial conditions in terms of $\mathbf{x}$-coordinates are

$$\mathbf{x} = \mathbf{r}/a \quad \text{and} \quad \dot{\mathbf{x}} = ((H - \dot{a}/a)\mathbf{r} + \delta\dot{\mathbf{r}})/a. \tag{31}$$

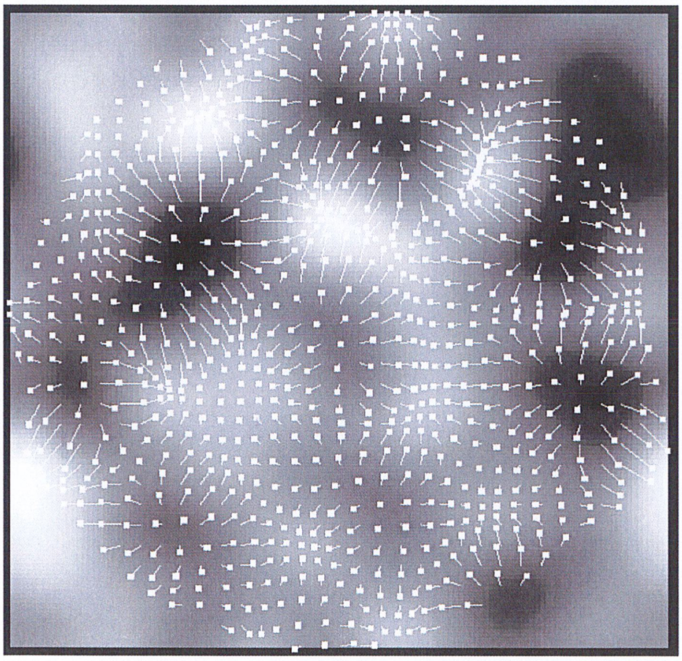


Figure 5: It was Dmitriev and Zeldovich who first wrote down the equations of motion for particle dynamics in re-scaled coordinates $\mathbf{x} = \mathbf{r}/a(t)$. They are precisely equivalent to the ordinary Newtonian equations (29) written in terms of physical coordinates $\mathbf{r}$. The 'scale factor' $a(t)$ has no dynamical significance; it is simply a book-keeping device and is completely arbitrary. However, by a judicious choice of $a(t)$ – choosing it to obey the Friedmann equation – we can arrange that the effective force term in (32) is caused only by the density fluctuation. If we start with particles on a grid within a sphere the force vanishes. If we perturb the particles off the grid as shown here – though in 2-dimensions rather than 3 – the force can be considered to be the sum of those caused by a set of little 'dipole' sources (a positive mass at the head and a negative mass at the tail of each of the little 'tadpoles'.

The sum in (30) will have two components: A 'zeroth order' acceleration that, in the limit that the grid spacing becomes very small, is the same as the gravitational acceleration of a uniform density sphere, which grows linearly with $\mathbf{x}_i$, plus a perturbation determined by the displacements from the grid (we may think of the source of the gravity being that of the unperturbed grid of particles plus that of a set of dipole sources). If we define the number density of particles in $\mathbf{x}$-space $n(\mathbf{x}) \equiv \sum_i \delta(\mathbf{x} - \mathbf{x}_i)$ and $\delta n \equiv n - \overline{n}$ with $\overline{n}$ the inverse of the grid cell volume in $\mathbf{x}$-space, equations (30) become

$$\ddot{\mathbf{x}}_i + 2\frac{\dot{a}}{a}\dot{\mathbf{x}}_i - \frac{Gm}{a^3}\int d^3x\, \delta n(\mathbf{x}) \frac{\mathbf{x} - \mathbf{x}_i}{|\mathbf{x} - \mathbf{x}_i|^3} = -\left(\frac{\ddot{a}}{a} + \frac{4\pi G m \overline{n}}{3a^3}\right)\mathbf{x}_i. \tag{32}$$

But since $a(t)$ is arbitrary we may *assert* that $a(t)$ is such that the RHS of (32) vanishes. Since $m\overline{n}$ is the mean mass density in $\mathbf{x}$-coordinates, that means that $m\overline{n}/a^3$ is the mean mass density $\overline{\rho}$ in $\mathbf{r}$-coordinates, so vanishing of the RHS is

$$\boxed{\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\overline{\rho}} \tag{33}$$

which is the Friedmann acceleration equation for a uniform density sphere (note that the density here is proportional to $1/a^3$ and so varies with time).

Defining the *density contrast* $\Delta(\mathbf{x}) \equiv \delta n(\mathbf{x})/\overline{n}$, vanishing of the left hand side implies

$$\boxed{\ddot{\mathbf{x}}_i + 2\frac{\dot{a}}{a}\dot{\mathbf{x}}_i = G\overline{\rho}(t)\int d^3x\ \Delta(\mathbf{x})\frac{\mathbf{x}-\mathbf{x}_i}{|\mathbf{x}-\mathbf{x}_i|^3}.} \tag{34}$$

These equations, first derived by Dmitriev and Zel'dovich are those that are solved in so-called 'N-body' simulations. An example is shown in figure 5 in which the initial conditions are taken to be those predicted to arise from quantum fluctuations of the inflaton field during an early inflationary phase. To be more precise, the 'background' Friedmann acceleration contains, in addition to the mean density of matter, the effect of the dark energy on the universal expansion. Also, the calculation of the force term – which, as written, would be expensive to evaluate, is computed in a clever manner, with the long-range force component obtained by solving Poisson's equation numerically, making use of fast Fourier transforms for efficiency. But the essential physics is contained in the -Z equations above.

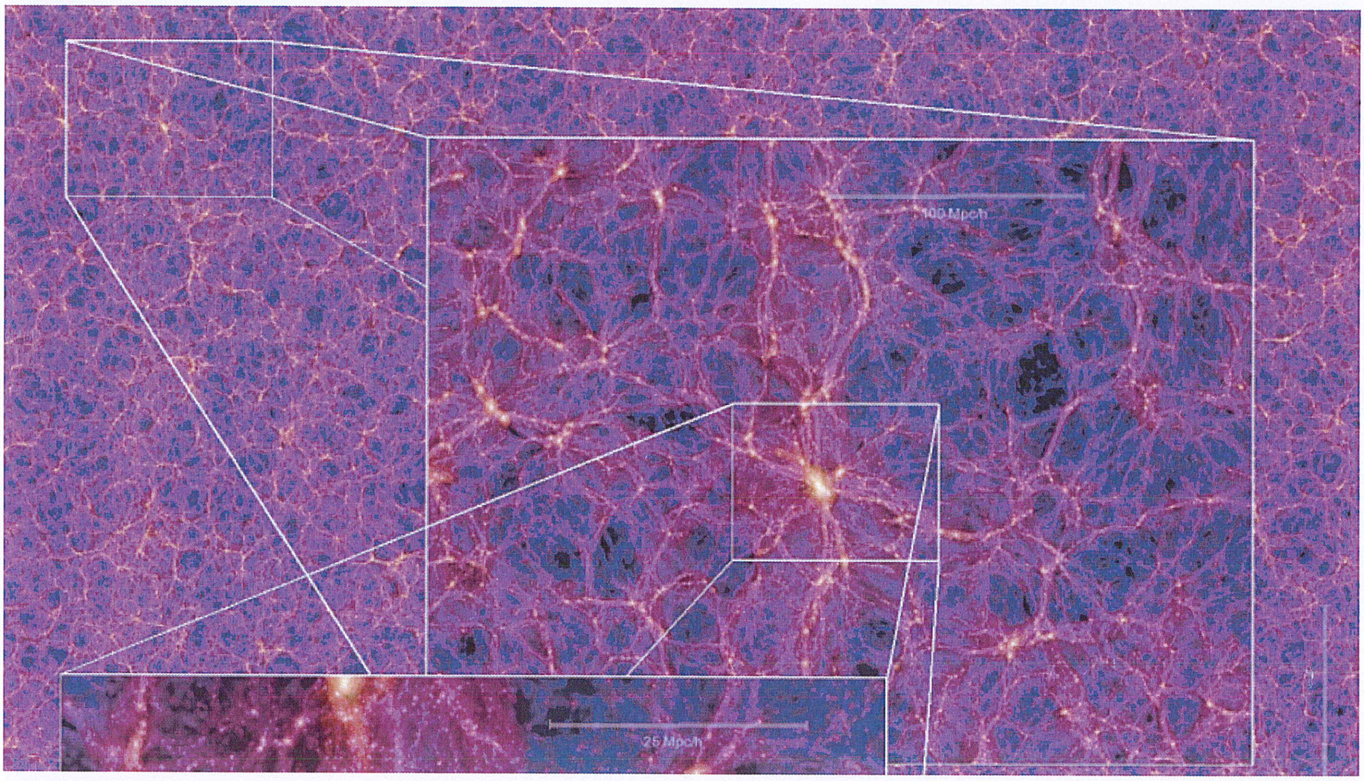


Figure 6: DM particles in the Millennium simulation

Comparing with the original expression of Newton's law (29) we see two changes. First, there is an additional *Hubble damping term* $2H\dot{\mathbf{x}}_i$. As it is proportional to the velocity $\dot{\mathbf{x}}$, it acts like a friction. We will see modifications like this in other equations of motions (Maxwell's equations for radiation, or the Klein-Gordon equations for a scalar field). While often described as being due to the gravitational field of the expanding universe – i.e. the effect of the 'expansion of space'. But in reality, such terms simply result from working in a non-physical coordinate system. Another useful way to think about the damping term is that, just as a photon will be seen to have energy that is decreasing by FOs that it is passing, massive particles have momenta that appear to decrease (again as measured by observers who are expanding away from each other). Secondly, we see that the right hand side – the so-called *peculiar* gravitational acceleration – is driven only by the matter density fluctuations $\delta\rho(\mathbf{x},t) = \overline{\rho}(t)\Delta(\mathbf{x},t)$.

6.2 Linerarised equations for structure growth

If we have a situation like that shown in figure 4, where the displacement of the particles is smooth and continuous, we can think of the displacement as a field $\delta\mathbf{x}(\mathbf{x})$ obeying

$$\ddot{\delta\mathbf{x}} + 2H\dot{\delta\mathbf{x}} = -\boldsymbol{\nabla}\phi \tag{35}$$

where the gradient operator is in $\mathbf{x}$-coordinates, and ϕ is the potential generated in Poisson's equation $\nabla^2\phi = 4\pi G\overline{\rho}\Delta$.

If the displacements are small compared to the length scale of the density fluctuations, the latter will be small in amplitude, and, working only to 1st order in the density perturbation Δ, this is just given by the divergence of the displacement field: $\Delta(\mathbf{x}) = \boldsymbol{\nabla}\cdot\delta\mathbf{x}$.

So if we take the divergence of the structure growth equations, these become

$$\boxed{\ddot{\Delta} + 2H\dot{\Delta} - 4\pi G\overline{\rho}\Delta = 0.} \tag{36}$$

These equations admit solutions which are the sum of growing and decaying modes: $\Delta(\mathbf{x},t) = \Delta_+(\mathbf{x})D_+(t) + \Delta_-(\mathbf{x})D_-(t)$.

In general, the time evolution functions must be obtained numerically. The situation is greatly simplified in the case of a critical density universe: $\Omega_{\mathrm{m}} = 1$ (and hence $\Omega_{\mathrm{k}} = 0$) for which $4\pi G\overline{\rho} = 3H^2/2$ and $H = 2/3t$ and, postulating a power-law $D \propto t^\alpha$, the differential equation becomes the quadratic one: $3\alpha^2 + \alpha - 2 = 0$, with solutions $\alpha = -1$ (the decaying mode) and $\alpha = 2/3$, for which $\Delta \propto t^{2/3} \propto a$.

For density perturbations introduced at very early times the structure will be essentially all in the growing mode. The way in which initially small density perturbations get amplified with time is historically dubbed *gravitational instability*. But that is something of a misnomer; what is happening here is that the structures are evolving in an expanding in such a way that the potential perturbations associated with the structures is actually constant in time.