Universe View through GWs

Ghanashyam Date
Chennai Mathematical Institute

This write-up is an expanded version of the webinar given in the Outreach lecture series of Kerala Theoretical Physics Initiative (KTPI) on September 28, 2024.

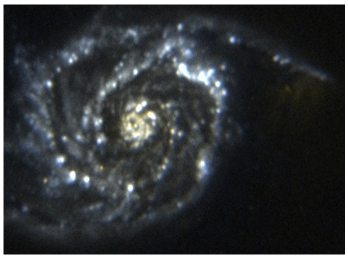
Contents

I. ELECTROMAGNETIC VIEW OF THE UNIVERSE

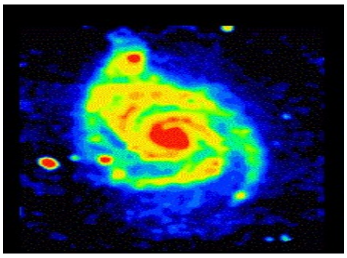
Look up in the sky and 'see' planets, stars, galaxy, . . . , star clusters, galaxy clusters. With different filters on, this is how the whirlpool galaxy looks:

infrared

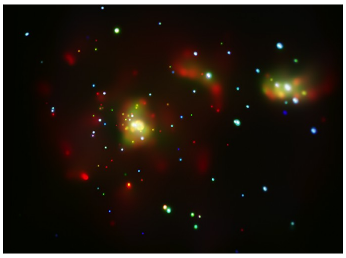
visible



ultraviolet



radio



x-ray

GW?

Credits:
http://www.physci.mc.maricopa.edu/Astronomy/astlabs/ast114/galaxy-lab/m51.htm
http://coolcosmos.ipac.caltech.edu/cosmic_classroom/multiwavelength_astronomy/multiwavelength_museum/m51.html

Note how different the view is! The views reveal that there are different types of stars, dust, . . . with different dominant processes producing EM waves in different frequency bands. The images show the distribution of these difference 'sources'.

How far are the different types of objects? This is a somewhat elaborate process, involving many stages constituting the "cosmological distance ladder". The distances to the nearby planets, stars are estimated by trigonometry and parallax method. Subsequent methods involve estimating the luminosity of the object and comparing it with the luminosity measured on earth and inferring the *luminosity distance* from the ratio using the inverse square law. The absolute luminosities are estimated by using the *variable stars* which have a definite relation between the period of variability of brightness and the absolute luminosity, using novae, supernovae, brightest galaxies etc. Here is a summary:

Nearby stars	Triangulation/parallax	~ 30 pc
Star clusters	main sequence photometry	$\sim 10^5$ pc
Galaxy	variable stars	$\sim 10^6$ pc
Inter-galaxy	Novae, brightest stars	$\sim 10^7$ pc
Galaxy clusters	supernovae, brightest galaxy	$\sim 10^{10}$ pc

The first non-trivial observation: If we look at only the objects which beyond about 200 Mpc, then their distribution is almost isotropic $\rightarrow$ *universe is isotropic on large scale.*

Invoke the *cosmological Principle* to assume that the same isotropic view will hold if viewed from any other galaxy. Thus, imagine the universe to be a stack of spatially homogeneous three dimensional slices.

The second non-trivial observation, the Hubble law: *Galaxies seem to be receding from us with a rate proportional to their distance from us.* $\rightarrow$ Universe is expanding!

What governs this expansion?

II. ENTER GENERAL RELATIVITY

Einstein's General relativity which a relativistic generalization of Newtonian gravity, provides a geometrical framework which naturally accommodates a spatially homogeneous expanding universe with specific equation governing the expansion.

Relativity – both special and general – suggest *new entities* in the cosmos. A white dwarf star with a instability beyond the Chandrasekhar limit of about 1.4 solar mass and a neutron star with a corresponding Tolman-Oppenheimer-Volkoff (TOV) limit of about 2.5 solar mass. And beyond these limits these stars collapse to *a black hole*. These are highly compact and heavy objects. White dwarfs have a radius of about 7000 km while Neutron stars have radii of about $10 - 15$ km.

And crucially, the general theory also predicts propagating gravitational effects – gravitational waves. These also carry energy, momentum, angular momentum and provide a new instability for bound gravitational systems eg two orbiting bodies heading towards a merger. How may these complement the electromagnetic view?

The quantitative estimates suggest that tightly bound binaries can actually merge and the best candidates are merger of neutron stars and black holes. When the gravitational wave astronomy develops further, we will see the distribution of such compact binaries and their remnants in the sky. These could not have been seen electro-magnetically (barring NS mergers). The blank box would get filled.

Of course such mergers could also have been happening throughout the history of the universe and such mergers may be expected to be distributed in a stochastic manner. Since

GWs can exist and propagate independent of matter, there could have be some unknown mechanism eg *quantum fluctuations* at the very beginning of the universe. These could also produce a stochastically distributed GW background. Together these two are expected to produce a *gravitational wave background*. The "sources" responsible for this background are appropriately thought of as *(spatially) non-compact sources*. By contrast, the merging systems of binaries, triples, . . . are *(spatially) compact sources*.

We have other examples of cosmic background, the *CMB* and possible but yet to be directly observed *Cosmic Neutrino Background*. Recall that CMB is produced when the plasma of protons, electrons and photons gets cooled and diluted enough so that photons don't have enough charged particles to scatter off. Photons decouple and free stream and constitute the background radiation. CMB is thus a snapshot of the surface of last scatter (LSS), approximately 3×10^5 years after the Big Bang. There is a corresponding decoupling of neutrinos at earlier time, about a second after the Big Bang and the free streaming neutrinos are expected to provide a snapshot of earlier time.

Gravitational waves couple so weakly with matter that they are essentially always decoupled and free streaming. Their background is thus expected to provide a snapshot of universe just after the Big Bang. We have no clue of what to expect for such a background.

A. But what are GWs? A very brief commentary

- Generically, waves are localized deviations/disturbances/patterns that propagate eg pulses. There could can be un-localized patterns that are time dependent eg plane waves or localized but non-propagating, time dependent patterns eg standing waves;

  Their space-time dependence is determined by some partial differential equation of a hyperbolic character. Such an equation can be linear (Maxwell in free space) or non-linear (fluids, GR, Yang-Mills etc).

- The most common way to study such class of solutions, is to *linearize* the equation about some *background* configuration. This is how we view ripples over a lake. Such solutions are perturbative in nature. Maxwell equations in free space are already linear and don't need any "background".

  Such a linearization was indeed done by Einstein himself for his equation with the Minkowski space-time as the background. He in fact derived the leading contribution to the radiation due to a time varying, spatially bounded matter distribution – the famous *quadrupole formula*.

- Initially, it was quite confusing to identify in an un-ambiguous manner "propagating solutions" - the propagation speed seemed to depend on coordinates chosen ("Gravitational waves propagate at the speed of thought" attributed to Arthur Eddington). It was also unclear if *these* gravitational waves i.e. propagating solutions of the *linearized* equation, are actually approximations to some solutions of the full non-linear equation. Given that there is no locally conserved stress tensor for gravity, it was unclear if there is any exact solution which can be interpreted as carrying energy, momentum, angular momentum etc. Somehow, one should identify "radiative solutions" (as hinted by the quadrupole formula) of the non-linear equation. This problem was solved by Bondi (1957), Bondi-Pirani-Robinson (1958-60) and Bondi-Burg-Metzner (1962) in a series of papers. Their inputs were:

  • There is some source which is spatially bounded and an asymptotic region can be identified; In this region, the vacuum equations hold;

  • With suitably adapted coordinates ("Bondi-Sachs coordinates"), a class of solutions satisfying the *"no incoming radiation" condition*, have a particular *asymptotic fall-off form* and are characterized by *mass-aspect function, angular momentum aspect and the news tensor*. Among these, a sub-class of solutions for which the *news function* (square of the news tensor) is non-zero which can be interpreted as radiation. This characterizes gravitational radiation; [*See the Azadeh Reference for asymptotic form.*]

  • Suitable quantities can be defined with the interpretation of the energy, momentum, angular momentum which change *iff* the news function is non-zero. Thus, radiation can be interpreted as conveying energy, momentum, angular momentum from the source region to far away. This analysis holds without any linearization and goes beyond the quadrupole formula. It also does not refer explicitly to any source dynamics.

  • Furthermore, linearized form of these equations ($\Lambda = 0$) have solutions with non-zero News function provided the provided the source multipoles beyond the dipole have accelerated time dependence.

  Thus, for the *spatially bounded sources*, at the level of non-linear equation, solutions exist which can be interpreted as radiation and the linearized solutions do provide an approximation.

- As an aside, we note that when Einstein derived the quadrupole formula, cosmological constant was not known and the equations did not have it. However, current cosmological model suggest that $\Lambda > 0$. Solutions of the linearized equation with a corresponding quadrupole formula have been derived. The formula contain Λ-dependent corrections which however are tiny for astrophysical, compact sources. A Bondi-Sachs type analysis for $\Lambda > 0$ has also been carried out. [*see Compère-Fiorucci- Ruzzicon,* *arxiv.org/pdf/1905.00971 and references therein.*]

- Not only does it provide a precise theoretical characterization of gravitational radiation, it also enables us to compute the *in-spiral of a binary orbit* in a quantitative way using the energy balance equation. This was superbly confirmed in the orbital decay of the Hulse-Taylor binary pulsar which constituted the first (indirect) evidence of gravitational waves.

  The subsequent development of post-Newtonian/Post-Minkowski framework of systematically going beyond the linearised solution provided a detailed in-spiral history *closer to merger* and the characteristic 'chirp' wave form.

  In summary: *GWs are indeed physical, have two transverse polarizations denoted as '$+'$ and '$\times'$, and propagate at the speed of light.*

B. Non-compact sources

What about sources which are not compact? These could arise from a stochastic distribution of compact sources all over *or* there could be some "primordial sources" i.e as yet unknown production mechanisms that are inherently non-compact. Such sources cannot be ascribed to any localized region. There are no asymptotic regions now and nothing to specify any fall-off conditions. Notion of *radiation* does not make much sense but we could still have *waves*. We have to seek guidance from a perturbative route and begin with linearization about some background which can support non-localized sources. The FLRW space-times are the natural choice The observed large scale isotropy also supports the perturbative approach. The perturbations can of course be spatially inhomogeneous and this manifests on our celestial sphere as anisotropies[1].

In either of the two ways to conceive of non-compact sources, we have nothing to go by about their distribution. So the only reasonable assumption that we can make is that such sources are generically distributed *inhomogeneously in a stochastic manner* and we can only hope to have information of a statistical nature. We will generically call such non-compact sources as *cosmological perturbations.*

We refer to gravitational waves produced by such sources as *Gravitational Wave Background*

[1] Just consider alternate dark and light vertical bands on a piece of paper and draw a circle at the center cutting the bands. The arcs will be dark and light showing an anisotropic distribution on the circle.

(GWB). The background produced by a distribution of compact sources is called *AstrophysicalGWB* while that due to primordial sources is called *StochasticGWB*.

How do we parametrise such backgrounds in a measurable manner?

We already have a dominantly isotropic distribution of matter with a small amount of anisotropy, caused by cosmological perturbations. Let us schematically denote the inhomogeneity caused by non-compact sources, in terms of a real valued stochastic variable $\delta\phi(t, \vec{x})$. Invoking only the cosmological principle (spatial homogeneity), we can write,

$$\delta\phi(t, \vec{x}) = \int \frac{d^3k}{(2\pi)^{3/2}} e^{-i\vec{k}\cdot\vec{x}} \left[z_{\vec{k}} \; u_k(t) + z^*_{-\vec{k}} \; u^*_k(t) \right] \; , \; k := |\vec{k}| \; . \quad (1)$$

Here, the $u_k(t)$ denote the complex solutions of the linear differential equation satisfied by $\delta\phi(t, \vec{x})$.

The stochastic nature of the inhomogeneity is stipulated as:

$$\langle \delta\phi(t, \vec{k}) \rangle = 0 \quad ; \quad \langle \delta\phi(t, \vec{k}) \delta\phi(t, \vec{l}) \rangle \; = \; \delta^3(\vec{k} - \vec{l}) \Delta^2_\phi(t, k) \quad \Leftrightarrow \quad (2)$$

$$\langle z_{\vec{k}} \rangle = 0 \quad ; \quad \langle z_{\vec{k}} z^*_{-\vec{l}} \rangle \; = \; \delta^3(\vec{k} - \vec{l}) \mathcal{A}(k) \quad , \quad \text{with} \quad (3)$$

$$\Delta^2_\phi(t, k) \; = \; 2\mathcal{A}(k) |u_k(t)|^2 \quad , \quad P_\phi(t) := \frac{k^3}{2\pi^2} \mathcal{A}(k) \; . \quad (4)$$

The $P_\phi(k)$ is called the *power spectrum* of the inhomogeneity $\delta\phi$. It is *dimensionless. If* all other correlation functions are determined in terms of this two point function, then the inhomogeneity is *Gaussian distributed* and all statistical properties of $\delta\phi(t, \vec{x})$ are specified entirely in terms of the power spectrum.

The primary observable for a stochastic background is the power spectrum which is to be determined observationally and any candidate theory of generation and evolution of perturbations should compute it.

We need to digress on the cosmological background to formulate perturbations and infer the deterministic evolution equations satisfied them.

III. COSMOLOGICAL BACKGROUND

Cosmological background space-times are premised on spatial homogeneity and the currently favored model is the *spatially flat* FLRW model. The universe is viewed as a stack of spatial hyper-surfaces, all orthogonal to the “fundamental observers” encoded by a time-like vector field $u^\mu(t,\vec{x})$. The coordinates are chosen such that $u^\mu = (1,\vec{0})$. The metric and the stress tensor take the form,

$$\Delta s^2 = -\Delta t^2 + a^2(t)\Big(\sum_{i=1}^{3}(\Delta x)_i^2\Big) \quad , \quad T_{\mu\nu} = \rho(t) u_\mu u_\nu + P(t)(u_\mu u_\nu + g_{\mu\nu}) \ . \qquad (5)$$

The Einstein equations, $R_{\mu\nu} - \frac{R}{2}g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}$ reduce to ($G = 1 = c$ units, overdot denoting $\frac{d}{dt}$),

$$\begin{aligned} 3\frac{\ddot{a}}{a} &= -4\pi(\rho + 3P) && \text{(Raychaudhuri equation);} \\ 3\left(\frac{\dot{a}}{a}\right)^2 &= 8\pi\rho && \text{(Friedmann equation);} \\ \dot{\rho} &= -3(\rho + P)\frac{\dot{a}}{a} && \text{(Conservation equation).} \end{aligned}$$

We need to specify a relation between the pressure and the energy density, called the equation of state. The three common choices and the corresponding solutions are given by,

$$\begin{aligned} P &= 0 \ , \ \ \rho a^3 = \text{constant} \ , \ \ a(t) = C_M t^{2/3} && \text{(Matter dominated);} \\ P &= \frac{1}{3}\rho \ , \ \ \rho a^4 = \text{constant} \ , \ \ a(t) = C_R\sqrt{t} && \text{(Radiation dominated);} \\ P &= -\rho = -\frac{\Lambda}{8\pi} \ , \ \ a(t) = C_{dS}e^{\sqrt{\Lambda/3}t} && \text{(de Sitter).} \end{aligned}$$

It is common to use the *conformal time*, η defined by $d\eta := a^{-1}dt$. It follows that η is *negative* for de Sitter, positive for the radiation/matter dominated epochs and *always increases with* t. The derivatives w.r.t. η are denoted by primes.

The first two equations imply that if $(\rho + 3P) > 0$, our expanding universe will continue to expand forever and therefore always shrinking as we view back in time. Given the current rate of expansion, $H_0 := \frac{\dot{a}}{a}\Big|_{now}$, the universe must have had zero size, a finite time ago (the Big Bang). The observed large scale isotropy suggests that the past light cones of the antipodal points must have intersected to establish such a correlation. Alas, this is not possible in the finite age of the universe estimated from the first two equations of state. This is the so-called *horizon problem*. It can be solved by *postulating an epoch of accelerated expansion or inflation*. The simplest model of this is the de Sitter epoch with a positive cosmological constant. More general models of inflation, minimally introduce a spatially homogeneous scalar matter field with a potential.

Here is a brief summary of the main epochs in the evolution of our continuously expanding universe:

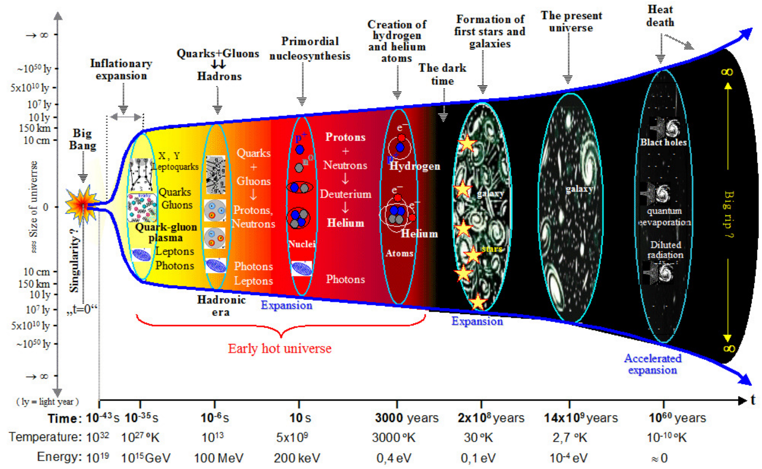
1. The big bang creation of the universe: The unknown era beyond classical GR. All scales are Planck scales - $10^{-35}m, 10^{-8}kg, 10^{-44}sec, 10^{19}GeV, 10^{32}\ ^{\circ}K$ etc.

2. Inflation: A brief period of exponential expansion (about 70 e-folds $\sim 10^{30}$ expansion factor), consists of gravity (metric) and scalar field (inflaton);

3. Re-heating: Decay of inflaton and transfer of energy to produce standard model particles;

4. Radiation dominated: Largely protons, electrons and photons till about the decoupling of photons at LSS.

5. Matter dominated: Dark ages followed by first stars and subsequent structure formation.



Credits: https://astronuclphysics.info/Gravitace5-4.htm

The always expanding universe provides a natural scale given by the inverse of the expansion rate: $d_H(t) := (H)^{-1} = (\dot{a}/a)^{-1}$, the "Hubble radius". This provides a time dependent, length scale for comparison. For matter/radiation dominated epochs, $a(t) \sim t^{\alpha}, 0 < \alpha < 1$, the Hubble radius $\propto t$ while for de Sitter, it remains constant at $\hat{H} := \sqrt{3/\Lambda}$. By contrast, all length scales eg wavelengths, sizes of objects etc scale with $a(t)$ i.e. increase as $\sim t^{\alpha}$ or $e^{\hat{H}t}$.

Hence the ratio of the Hubble radius and any physical scale *increases* as $t^{1-\alpha}$ during deceleration and *decreases* $e^{-\hat{H}\ t}$ during de Sitter inflation (acceleration). This has important implications for the evolution of the amplitudes of inhomogeneities.

Our presumption has been that perturbations in the matter stress tensor and the geometry cause the inhomogeneities. How do we know such perturbations exist?

A. Cosmological anisotropies or "reality of perturbations"

The background quantities are homogeneous and isotropic as motivated by the observations of the distribution of galaxies. This is reflected also in the CMB angular distribution, which is beautifully isotropic. However, there are anisotropies reflecting the inhomogeneities in the plasma producing the CMB. These inhomogeneities in turn are thought to be produced by perturbation in the epochs preceding the CMB.

Two noteworthy properties of the observed CMB (temperature) anisotropies are: (a) The amplitudes of the anisotropies $\Delta T/\bar{T}(\hat{n})$ are about 10^{-5} times smaller than the those of the isotopic CMB. As per COBE normalization of the power spectrum $P_S(k) := A_S(k/k_*)^{n_S-1}$ the amplitude $A_S \sim 2\times10^{-9}$ and the corresponding one for tensor perturbations is $0.036A_S$; (b) the angular distribution is resolvable at about 10 arc-minutes (Planck Satellite). 10 arc-minutes corresponds to about $\ell = 2160$ and since the LSS has a radius of about 40 billion light years[2], the linear scale corresponding to 10 arc-minutes is about 10^8 light years or about 30 Mpc. The largest possible length scale, angular separation of π corresponds to about $36Gpc$. The present Hubble-radius is about 4 Gpc.

The CMB observations *confirm* the existence of perturbations in a wave length/frequency band along with the size of these perturbations.

There is nothing to preclude perturbations outside the scales seen in the CMB anisotropies. Can these be revealed in other observations? The answer is 'yes' and below is an indicative summary of which scales can be seen in which wavelength/frequency bands.

Experiment	Sensitivity window (f)	wavelength range
Ground based interferometer	$1 - 10^3 Hz$	$300 - 3\times10^5\ km$
Space based interferometers	$10^{-5} - 0.1Hz$	$10^6 - 10^{10}\ km$
Pulsar Timing Arrays	$3\times10^{-9} - 10^{-6}Hz$	$10^{-1} - 30\ lyr$
CMB	$3.4\times10^{-19} - 7\times10^{-18}Hz$	$4 - 10^2\ Gpc$

While the goal of the interferometric observations is direct detection of waveforms from compact sources, a set of such detectors can also be used to detect the GW backgrounds i.e. estimate the power spectrum in the corresponding frequency bands.

[2] The LSS radius is the η elapsed times the scale factor now in the $c = 1$ units. Taking $a_{now} = 1$, the epoch as matter dominated and neglecting the time interval from the big bang to decoupling, we get $\eta_{now} \approx \int^{t_0} \frac{dt}{(t/t_0)^{2/3}} = 3t_0 \approx 40\times10^9$ years.

Having established he existence of perturbations, we take a look at evolution of cosmological perturbations, limiting ourselves to gravitational perturbations alone.

IV. GRAVITATIONAL PERTURBATIONS AND THEIR EVOLUTION

We begin by writing the perturbations as, $g_{\mu\nu} = \bar{g}_{\mu\nu} + \epsilon h_{\mu\nu}$ for the metric, $T_{\mu\nu} = \bar{T}_{\mu\nu} + \epsilon T^{(1)}_{\mu\nu}$ for the stress tensor and substituting in the Einstein equation.

It is important to recall from GR that *neither the coordinates x^μ nor the metric components $g_{\mu\nu}(x)$ are physically meaningful, only the $\Delta s^2 = g_{\mu\nu}\Delta x^\mu \Delta x^\nu$ are the basic physical quantities.* This in particular implies that any small difference between two metrics, $\delta g_{\mu\nu} \sim \epsilon(\bar{\nabla}_\mu \xi_\nu + \bar{\nabla}_\nu \xi_\mu)$, can be generated by a small change in coordinates, $\delta x^\mu = \epsilon \xi^\mu$, without changing Δs^2. Such $h_{\mu\nu}$ thus *do not represent a physical change.* Similarly changes in arbitrary tensors which are of the form, $\delta T \sim \epsilon \mathcal{L}_\xi T$, merely reflect coordinate change. It is conventional to call such unphysical changes as *gauge transformations.* The physically meaningful quantities are *gauge invariant.*

One may either use the manifestly gauge invariant quantities *or* fix the coordinates completely by imposing appropriate condition on the perturbation which directly represent physical quantities[3]. Both methods are used in practice.

For our discussion, we focus only on the fully gauge fixed form of the metric perturbations i.e. the gravitational ripples on the cosmological background, also called *tensor perturbations.* We also restrict to the *linear order* in perturbations. This means that we look for solutions of the equations,

$$\bar{\Box} h_{ij} = 0 \ , \ \bar{\nabla}_j h^j{}_i = 0 = \bar{g}^{ij} h_{ij} \ ; \tag{6}$$

The $\bar{\Box}$ is the d'alembertian of the background metric and the i, j are spatial indices. For primordial fluctuations, there are no matter sources while for perturbations evolving through subsequent epochs would typically have transient sources which are neglected in our discussion[4].

[3] In electromagnetism described in terms of $A_\mu(x)$, one may work with the $\vec{E}, \vec{B}$ (or $F_{\mu\nu}$) and functions thereof which are manifestly gauge invariant *or* use the A_μ which further satisfies the conditions such as $\partial_\mu A^\mu = 0 = A_0$ leaving only two independent variable reflecting the two polarizations.

[4] These statements hold at linear order (*primary primordial*). We can go to higher orders corrections which are sourced by the lower order ones. These are called *secondary, tertiary ... primordial* gravitational waves. We focus on the primary ones.

Given that the cosmological background has spatial homogeneity, we can Fourier decomposition of the perturbations and label the modes by $\vec{k}$. To linear order, the equations for each Fourier mode decouple. Their time dependence is determined by the evolving background, $a(t)$. The general solution $h_{ij}(\eta, \vec{x})$ can thus be developed as,

$$h_{ij}(t,\vec{x}) = \sum_{\sigma=+,\times} \int \frac{d^3k}{(2\pi)^{3/2}} e^{-i\vec{k}\cdot\vec{x}} [h_\sigma(\vec{k},t)\varepsilon^\sigma_{ij}(\vec{k})] \quad , \quad k^j e^\sigma_{ji} = 0 = \delta^{ij} e^\sigma_{ij} \ . \tag{7}$$

The $\varepsilon^\sigma_{ij}(\vec{k})$ denote the polarization tensors for a wave with wave vector $\vec{k}$ and the $h_\sigma(\vec{k},t)$ denote the corresponding amplitude.

In terms of $H_{ij}(\vec{k},\eta) := a(\eta)h_{ij}(\vec{k},\eta)$ (and $H_\sigma(\vec{k},\eta) := a(\eta)h_\sigma(\vec{k},\eta)$), the perturbations satisfy ,

$$H''_{ij}(\vec{k},\eta) + \Big(k^2 - \frac{a''}{a}\Big) H_{ij}(\vec{k},\eta) = 0 \quad \leftrightarrow \quad H''_\sigma(\vec{k},\eta) + \Big(k^2 - \frac{a''}{a}\Big) H_\sigma(\vec{k},\eta) = 0 \ .$$

For the three common epochs, $a(\eta) := a(t(\eta)) \sim \eta^\alpha$ with $\alpha = 2$ for matter dominated, $\alpha = 1$ for radiation dominated and $\alpha = -1$ for the de Sitter epochs. In all three cases,

$$\frac{a''}{a} = \frac{\alpha-1}{\alpha}\left(\frac{a'}{a}\right)^2 = \frac{\alpha-1}{\alpha}(aH(t))^2 \ \geq 0 \quad \Rightarrow (k^2 - a''/a) \lesseqgtr 0 \ . \tag{8}$$

Notice that in all three cases, $aH(t(\eta)) = a'/a \sim \eta^{-1}$.

The perturbation equation shows that for any given η, the modes with label k fall into two groups:

$$\begin{aligned} \frac{k^2}{a^2} &\ll H^2 \ \leftrightarrow \ (k\eta)^2 \ll 1 \ \leftrightarrow \ \lambda_{phy}(t) \gg d_H(t) \ \ \text{(super-Hubble)} \\ \frac{k^2}{a^2} &\gg H^2 \ \leftrightarrow \ (k\eta)^2 \gg 1 \ \leftrightarrow \ \lambda_{phy}(t) \ll d_H(t) \ \ \text{(sub-Hubble)} \end{aligned}$$

The super-Hubble modes $h_\sigma(k|\eta| \ll 1, \eta)$ loose the η-dependence and their amplitude remains constant while the sub-Hubble modes evolve as $h_\sigma(k,\eta) \sim a^{-1}(\eta)$ decreasing their amplitude. Since η^2 decreases with t during inflation, a super-Hubble modes remains super-Hubble through inflation. Likewise sub-Hubble modes remain sub-Hubble during late time evolutions. In particular, at the time of decoupling (LSS), the modes corresponding to $\ell \lesssim 20$ are super-Hubble and have retained their amplitude from the early epochs while the shorter wavelength modes being sub-Hubble their amplitude A_S will continue to fall as $a^{-2}(t)$. Some of the super-Hub modes can turn sub-Hubble at a later time and begin decreasing their amplitude.

As an explicit example, consider the evolution in a de Sitter epoch.

V. h_σ DURING A DE SITTER EPOCH

The general solution in these three cases is given in terms of spherical Bessel functions. In the de Sitter epoch, $\eta_{infl} \leq \eta \leq \eta_{end}$, the general solution takes the form,

$$\begin{aligned} h_\sigma(\eta, \vec{k}) &= z_\sigma(\vec{k})\ e_k(\eta) + z^*_\sigma(-\vec{k}) e^*_k(\eta)\ ,\ \text{(reality of } h_\sigma(\eta, \vec{x}) \text{ is used);} \\ e_k(\eta) &= \frac{\alpha_k}{a(\eta)}\Big(1 - \frac{i}{k\eta}\Big)e^{-ik\eta}\ ,\ \alpha_k \text{ are normalization constants.} \end{aligned}$$

The two free constants, $z_\sigma(\vec{k}), z^*_\sigma(-\vec{k})$ signify initial conditions and their distribution is stochastic. Suppress the polarization index σ on the $z_\sigma(\vec{k})$. We can consider three different cases: $h_\sigma(\eta, \vec{k})$ is (i) a classical stochastic function, (ii) is a classical stochastic field, and (iii) is a quantum field.

Stochastic function: $\langle z_{\vec{k}}\rangle = 0 = \langle z^*_{\vec{k}}\rangle\ , \langle z_{\vec{k}} z^*_{-l}\rangle := \mathcal{A}(\vec{k})\delta^3(\vec{k} - \vec{l})\ ,\ \mathcal{A}(\vec{k})$ is some unknown, positive function of $\vec{k}$ and we assume it to depend only on $k := |\vec{k}|$ to encode isotropy. The angle brackets denote an average w.r.t. the unknown probability distribution for the constants. Using these assumptions together with translational invariance and isotropy, we get the *autocorrelation function* as,

$$\begin{aligned} \langle h_\sigma(\eta, \vec{x}) h_\sigma(\eta, \vec{y})\rangle &= \int d^3k \frac{e^{i\vec{k}\cdot(\vec{x}-\vec{y})}}{(2\pi)^3}\ [\mathcal{A}(k)\ 2\ |e_k(\eta)|^2] \\ \therefore \langle h_\sigma(\eta, \vec{x}) h_\sigma(\eta, \vec{x})\rangle &= \int \frac{d^3k}{(2\pi)^3}\ [\mathcal{A}(k)\ 2\ |e_k(\eta)|^2] \ =\ \int_0^\infty \frac{dk}{k}(2)\left[\frac{k^3}{2\pi^2}\mathcal{A}(k)|e_k(\eta)|^2\right] \\ \text{with } |e_k(\eta)|^2 &= \frac{|\alpha_{\vec{k}}|^2}{a^2(\eta)}\Big(1 + \frac{1}{k^2\eta^2}\Big) \end{aligned}$$

The power spectrum is determined by the square bracket. The factor of 2 is due to the two terms contributing to the $\langle zz^*\rangle$. Both $\alpha_{\vec{k}}$ and $\mathcal{A}(k)$ are unknown. Notice that $a^{-2}(\eta) = H^2\eta^2$ for de Sitter and $[\cdots] \to [\frac{k^3}{2\pi^2}\mathcal{A}(k)(H|\alpha_k|/k)^2(k^2\eta^2 + 1)]$.

Stochastic classical field: The extra attribute requires the perturbation to satisfy the Poisson bracket relation. This translates into the requirement that $\{z_{\vec{k}}, z^*_{\vec{l}}\} = -i\delta_{\vec{k},\vec{l}}$. This additional condition involves the Wronskian of the solutions $e_k(\eta)$ and determines $|\alpha_k| = \frac{1}{\sqrt{2k}}$. The unknown function $\mathcal{A}(k)$ remains to be determined. The square bracket becomes, $[\cdots] = H^2(\mathcal{A}(k)/k^3)(1 + k^2\eta^2)$.

Quantum Field: The Poisson bracket condition becomes commutator and retains the determination of α_k. But now $\langle z_{\vec{k}} z^*_{\vec{l}}\rangle \to Tr[\hat{\rho}\hat{z}_{\vec{k}}\hat{z}^\dagger_{\vec{l}}]$. The density operator $\hat{\rho}$ though unknown, provides a handle for theoretical proposals for $\mathcal{A}(k)$.

In particular, for $\rho = |0\rangle\langle 0|$, the average becomes 1, the factor of 2 becomes (1) and $\mathcal{A}(k) = 1\ \forall k$ and the power spectrum becomes,

$$P_h(k,\eta) := \frac{k^3}{2\pi^2}|e_k(\eta)|^2 \ = \ \frac{H^2}{4\pi^2}(1 + k^2\eta^2) \ \rightarrow \ \frac{H^2}{4\pi^2} \text{ for super-Hubble } k\ . \tag{9}$$

Note that the determination of the power spectrum of the gravitational background is still through the CMB anisotropies observed through electromagnetic waves.

VI. ASTROPHYSICAL GWB

In the late universe, most of the primordial modes have been sub-Hubble and their amplitudes are expected to have decayed to negligible level. The relevant sources are compact sources with unknown distribution. These constitute the astrophysical GW background, also stochastic since the distribution of sources such are mergers of supermassive black holes throughout the late universe is not known. How do we determine their power spectra?

Here the main tool is the *Pulsar Timing Arrays.*

Pulsars: These are rotating neutron stars which beam the electromagnetic radiation (radio waves $3 - 10^9$ Hz) due to the rotating magnetic field anchored to the NS. Their rotation periods, range from milliseconds to a few seconds. These are extremely stable for millisecond pulsars with a fractional change in rotational frequency to be about $\delta\nu/\nu \sim 10^{-15}$. The radio pulses emanating from a pulsar suffers various *time delays* such as gravitational time dilation in Pulsar neighborhood, traversal in interstellar medium, delays due to intervening massive bodies, variation in receiver in orbit around the sun etc (Einstein, Romer, Shapiro, . . .). The times of arrival of pulses can be inferred and constitute the basic observation. Given the location of a pulsar and information about its local environment, interstellar medium etc one can estimate the time of arrival of pulses and compare it with the recorded TOA data. The *difference between the two*, called residual time delay is potentially due to GWB. This is estimated as follows.

Let a pulsar's sky location be in the direction $\hat{p}$. Let the infinitesimal distance along the line of sight be $\Delta s^2 - \Delta t^2 + \{\ (\delta_{ij} + h_{ij})\hat{p}^i\hat{p}^j\ \}d\alpha^2$. Since $\Delta s^2 = 0$ for light, we get $\Delta t^2 = (1 + h_{ij}\hat{p}^i\hat{p}^j)\Delta\alpha^2$. Integrating along the path connecting the pulsar to observer gives the elapsed time between emission and detection of the radio wave. Since the GW has t-dependence, the elapsed time for consecutive pulses differs from the pulsar period. If T denotes the pulsar period, then

$$z(t) := \frac{\Delta T}{T}(t) \;\approx\; \frac{\hat{p}^i\hat{p}^j\varepsilon^{\sigma}_{ij}(\hat{k})}{2(1+\hat{k}\cdot\hat{p})}h_{\sigma}(t)\ . \tag{10}$$

Here the gravitational wave has polarization σ, is moving along $\hat{k}$ and t is the time of arrival [*Maggiore, Vol II, Chap 23*].

Since the gravitational wave is supposed to constitute a stochastic background, we compute the pair-wise correlation $\langle z_a(t)z_b(t)\rangle$. Writing the hh correlation in terms of a power spectrum and carrying out the angular averaging over $\hat{k}$, we write

$$\langle z_a z_b\rangle \;=\; \left\{\int\frac{dk}{k}\left[\frac{k^3}{2\pi^2}\mathcal{A}(k)\right]\right\}\left[4\pi\mu_{Hellings-Downs}(\gamma_{ab})\right]\ ,\ cos\gamma_{ab} := \hat{p}_a\cdot\hat{p}_b; \tag{11}$$

$$\mu_{H-D}(\gamma) \;=\; \frac{1}{4}+\frac{1}{12}cos(\gamma)+\frac{1}{2}\left\{(1-cos\gamma)ln\left(\frac{1-cos\gamma}{2}\right)\right\} \tag{12}$$

The angular dependence is factored out and can be identified in the measurement of the correlation function.

In June 2023, the NANOgrav collaboration announced an evidence for GW background after analysis of 15 year data of 68 millisecond pulsars.

In Summary:

Waves are messengers from processes taking place far away. Gravitational waves fill-in where EM waves step aside. They do so in the form of *chirps*, as a background *hum* and also *leave imprints* on the polarization anisotropies of CMB which we did not discuss.

Here are some of the recent readable accounts.

Valerie Domcke, TASI lectures, arxiv:2409.08956.
Azadeh Maleknejad, MPA Lectures on GWs in Cosmology.
Caprini and Figueroa, CGWB, arxiv:1801.04268.
