L3 Astro - Section 4 - The Interstellar Medium

Nick Kaiser

October 27, 2020

Contents

List of Figures

1 The ISM: what is it? where is it found? what is its physical state?

1.1 An inventory of the ISM

Interstellar medium: everything in the galaxy that is outside of stars and planets

- atomic and molecular gas
  - emission and absorption – optical/IR and radio (21cm)
- dust
  - thermal emission (IR) – from satellites
  - scattering – absorption: spectral features – reveals chemical composition
- ionized gas (plasma)
  - pulsar dispersion – measures electron density along line of sight
  - interstellar scintillation (various types)
- hot plasma ($T \sim 10^6$K)
  - 'free-free' emission ⇒ X-rays
  - cyclotron radiation ⇒ radio emission
- cosmic rays
  - discovered with photographic emulsions in balloons
  - Cerenkov (Heaviside) arrays
  - created by acceleration in shocked magnetised matter
- magnetic fields
  - Zeeman splitting of atomic lines (solar corona)
  - Faraday rotation – measures line of sight **B**-field times plasma density

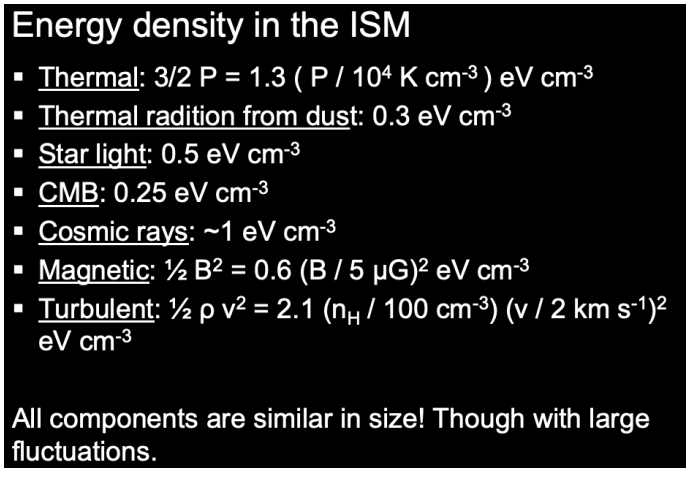


Figure 1: Energy densities of various constituents of the ISM. Give or take factors of a few, all the components of the ISM have broadly similar energy density $\sim$ 1eV/cm^3 as does, it turns out, starlight and the energy in the cosmic microwave background (CMB). The fact that there is rough 'equipartition' of energy strongly suggests that the different components are coupled through some kind of feedback process.

1.2 Maps of the ISM

- 3D maps:
  - neutral hydrogen (HI) radio emission – requires rotation curve ⇒ spiral arms (see below)
  - dust absorption – from locus of stars with known distances in colour-colour space
    - gives column density vs. distance

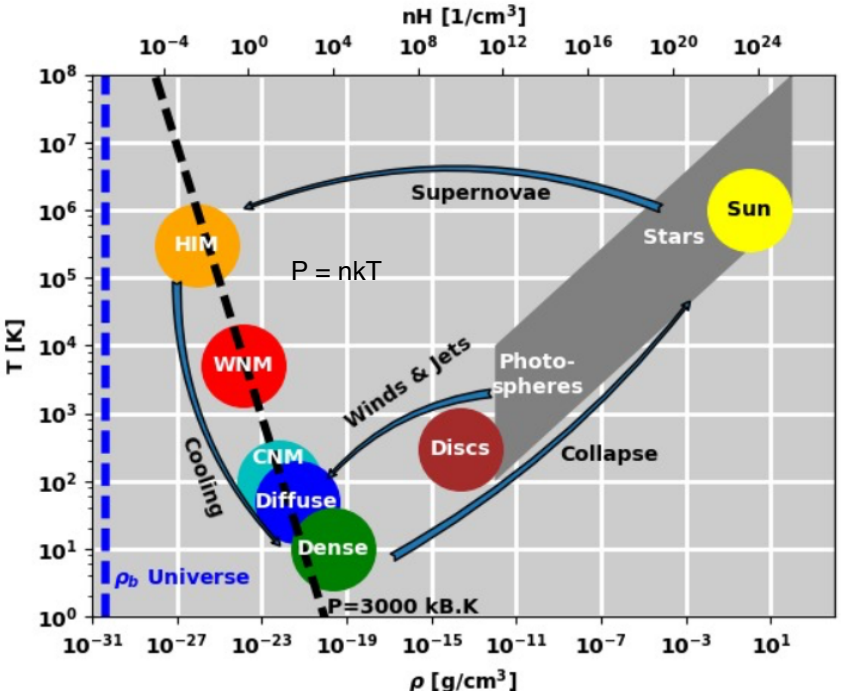


Figure 2: So called 'phases' of the ISM (courtesy Pierre Lesaffre) are shown on the left on a temperature-density ($T - n$) plot. On the right is shown the material in stars and their photospheres. The left-pointing arrows indicate feedback from star formation to the ISM from supernovae, winds and jets. The right pointing arrows show how gas (which starts off as hot interstellar medium (HIM)) can cool and become dense and ultimately form the stars.

    - derivative gives the space density

- 2D maps:
  - molecular clouds – sub-mm radio – confined to the disk
  - Planck satellite: multiple bands sensitive to different components

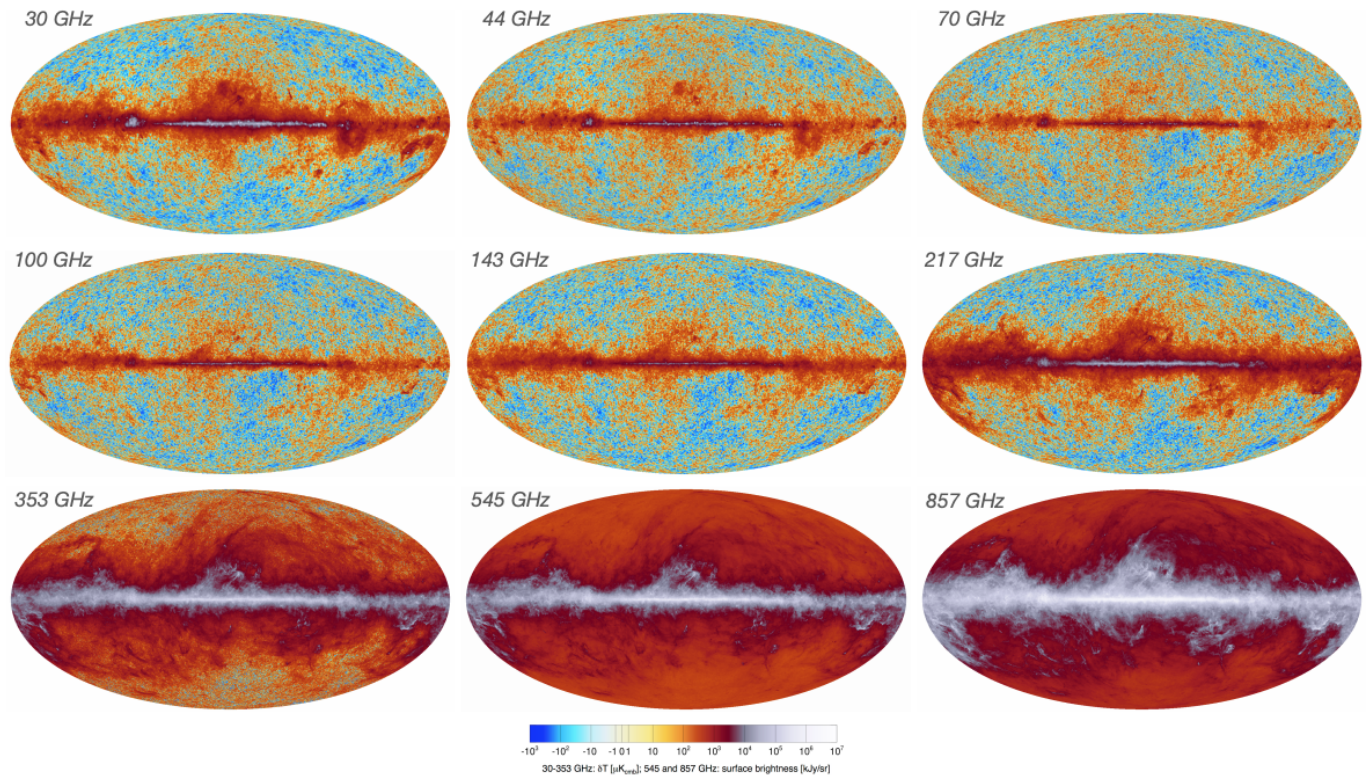


Figure 3: Maps of the sky from the Planck satellite in 9 pass-bands. These were made to measure the primordial temperature fluctuations of the cosmic microwave background (CMB). but also provide copious information on galactic 'foregrounds'. The higher frequencies show dust emission and the lower frequencies show radio synchrotron emission. Also present is emission from molecular clouds in the galactic plane.

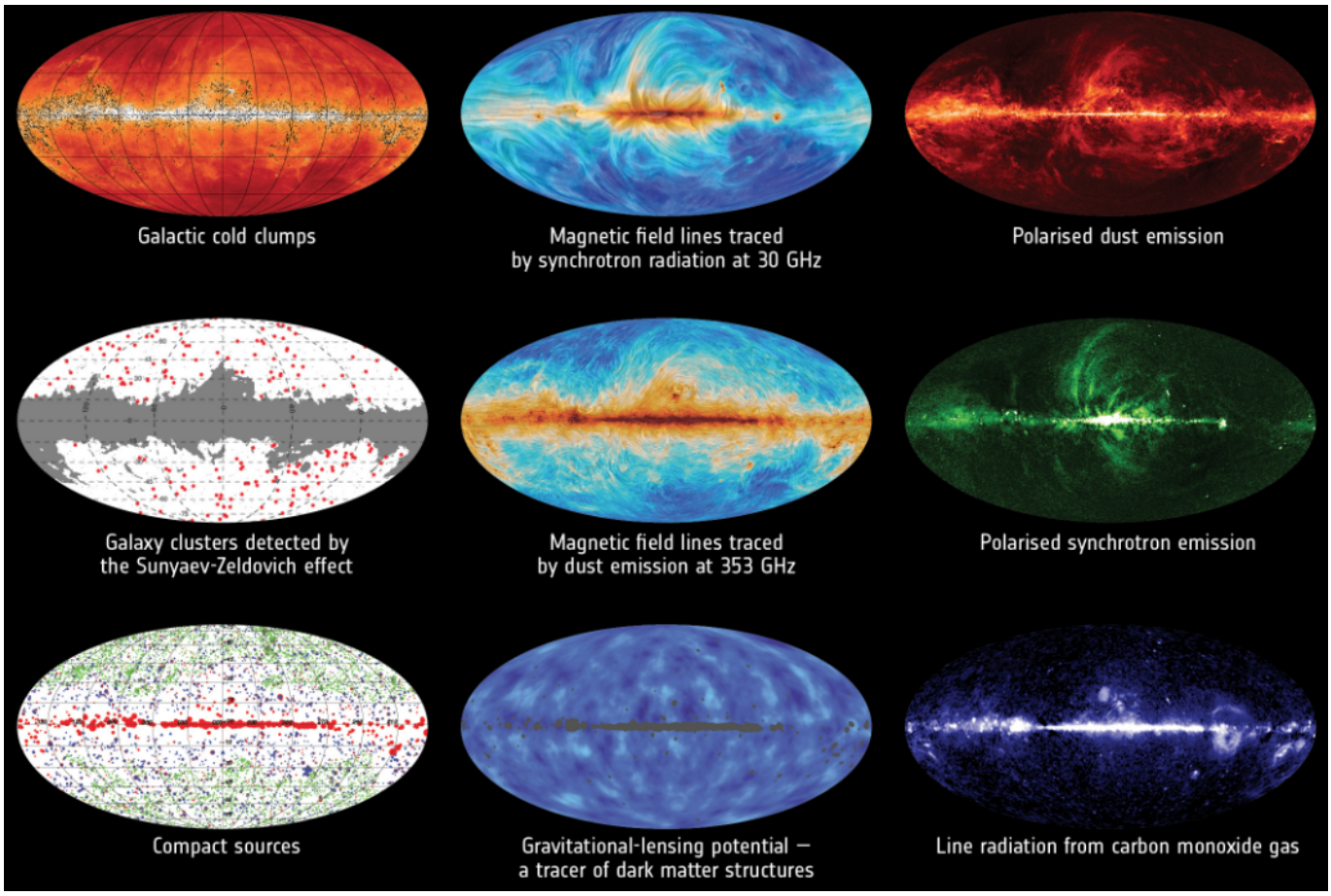


Figure 4: By making linear combinations of the intensity at different frequencies (and using polarisation information) the Planck team were able to construct maps of the various 'foregrounds'. That is everything coming from redshift less than 1000, and including several contributions from the ISM in our galaxy.

2 The 'two-phase medium' model

2.1 The 'phases' of the ISM

Various probes indicate that gas & plasma is found in 3 'phases': Cold dense clouds with $T \sim 10^2$K; hot gas at $T \sim 10^4$K; very hot X-ray emitting fully ionized plasma at $T \sim 10^6$K.

Broadly speaking these phases lie on a line in log-temperature - log-density space $nT \sim$ constant. Why? And why is the gas found in these different phases?

Name	State of H	Typical n (cm^{-3})	T (K) (gas)	Heating	Cooling	How observed	Remarks
Molecular clouds	H_2	>1000	10 - 80	Cosmic rays	CO, Far-IR from dust	CO (115 Ghz)	Dust has icy mantles
H I clouds	H	30	100	Photo-electrons from dust	[C II] (158 μm)	21-cm (emission, absorp)	"Diffuse ISM"
Warm H I	H	0.1	8000	"	"	21-cm emission	"
Warm H II	H^+	0.03	10^4	Photo-ionization of H	[O II], [S II]	Hα,[S II], nebular lines	Very faint but ubiquitous
Hot ISM	H^+	10^{-3}	$10^{6.5}$	Shocks from SNe	X-rays	X-rays	Little mass
H II regions	H^+	>100	10^4	H photo-ionization	[O III]	Hα, radio, other lines	Expanding, transient
Super-nova remants	H^+	(varies)	10^7	Shocks	X-rays, IR from dust	Optical, X-rays, IRAS	Dynamic!

Figure 5: Phases of the interstellar medium (from John Mathis's review). The stable phases (not the last two) consist of cold, warm and hot phases with distinctly different temperatures ($\sim$ 100K, 10^4K and $10^{6.5}$K respectively). They are in pressure equilibrium with one another, so they lie on a line of slope -1 on the $\log n - \log T$ plane ($nk_{\rm B}T \simeq$ constant) though there is, for example, a wide range of densities within molecular clouds, and there may be other forms of pressure support. They are subject to various types of heating mechanisms and also suffer energy losses (cooling). The existences of the different phases was first explained in the 50's by George Field in what is called the *'two-phase'* model for the ISM.

2.2 George Field's '2-phase' thermal instability model

Consider a roughly homogeneous medium consisting of gas being heated by sources (of radiation and maybe cosmic rays) and losing energy by collisional excitation followed by emission.

What happens if there is a region which is a bit denser than average?

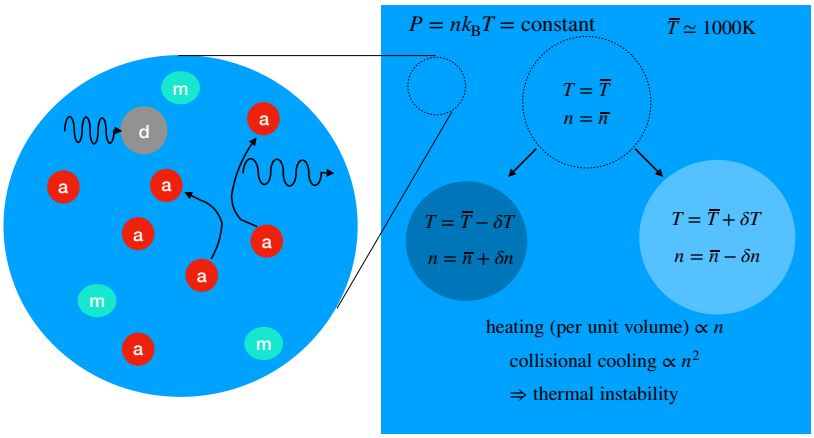


Figure 6: George Field's 2-phase thermal instability. Circle at left shows the constituents of the ISM (atoms, molecules and dust) and the processes: heating by an external source of photons and 'collisional cooling': an atom being excited in a collision and then radiatively de-exciting. At temperatures between $T \sim$ 100K and $T \sim 10,000$K this is unstable: if a region is compressed (expanded) a little it will radiate more (less) efficiently and still get cooler (hotter). This is a runaway process.

- It will radiate more (per unit mass) because 'cooling' $\propto n$.
  - note: collisions per unit *volume* $\propto n^2$
- But it will still receive the same energy input per unit mass because 'heating' is independent of n
  - provided m.f.p. of photons $\gg$ size of region
- So the temperature will drop.
- But hydrostatic equilibrium implies pressure balance.

- But pressure $P = nkT$, so the density n increases still further

This suggests a runaway instability. Any over-dense region will get cooler and even denser and any under-dense region will get hotter and even less dense.

Q: What stops this?

- The answer relies on the fact that the 'cooling rate' is temperature dependent
  - this depends on atomic and molecular physics
- At very low temperatures the energy of collisions is insufficient to excite molecules (as we have assumed)
  - so these stay in their ground states and the collisions will be elastic
- so the collisional cooling 'switches off' if the temperature falls below ~ 100K
  - and the dense gas stabilises in the 'cold-phase'
  - in molecular clouds
- Similar considerations show that the cooling rate from atomic processes increases rapidly around $T \sim 10^4$K.
  - and the gas becomes stable in the 'warm-phase'
  - the warm ISM

This was worked out by George Field and gives us his celebrated *two-phase* model for the ISM.

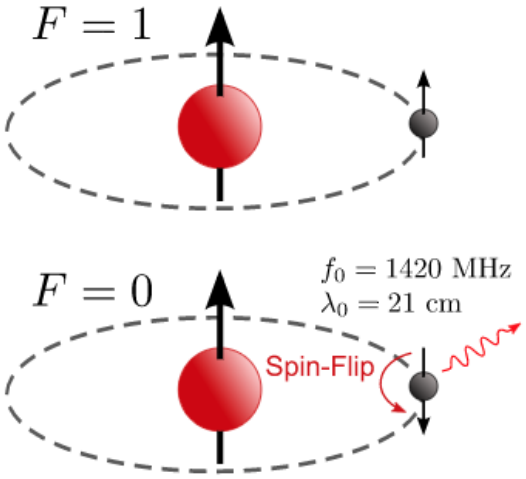
- If we had matter at intermediate temperature $100\text{K} \ll T \ll 10^4\text{K}$ then the instability would operate.
- Over-dense regions would radiate more effiently and cool down
  - but would stabilise at $T \sim 100$K
- Under-dense regions would radiate less efficiently and heat up
  - but will stabilise at $T \sim 10^4$K

Implications/predictions:

- Hydrostatic equilibrium: $\rightarrow$ pressure balance
  - Matter constrained to lie on a line of constant nT in the density-temperature $(n - T)$ plane
  - since $p = nkT$
- explains why we don't see matter at $100\text{K} \ll T \ll 10^4\text{K}$
- instead we see cold dense clouds with $T \sim 100$K embedded in a diffuse hotter medium
- matter will tend to be found on regions of the $p =$ constant locus in $n - T$ space where cooling varies rapidly with T
- similar considerations imply that gas with $T \gg 10^4$K would be unstable
  - under-dense regions would be unstable to 'run-away' to even higher temperatures
  - this is limited by the fact that the potential well of the galaxy can only contain plasma with $T \lessapprox 10^6$K
  - this is called the 'three-phase model'

Aside: cooling of gas is a crucial element of galaxy formation theory.

3 Mapping the MW with neutral hydrogen 21cm emission



- Electron and protons have spin → dipole magnetic moment.
- Hamiltonian (energy) depends on the relative orientation of the spins.
- Gives *hyperfine* transitions: ΔE is energy difference between parallel and anti-parallel configurations.
  - 'spin-flip' transitions → emission and absorption of λ =21 cm (radio) photons
  - it is a 'forbidden' line: transition rate $10^{-7}\mathrm{yr}^{-1}$
- It is extremely useful for mapping the structure of the Milky Way and beyond

3.1 21cm mapping of the galaxy

- External galaxies can be broadly split into *elliptical* and *spiral* morphologies
  - ellipticals are *pressure supported*
    - i.e. random motions of stars prevent collapse under gravity
  - spirals are flattened and *rotationally supported*
    - typically $\sim 100 - 300$km/s rotation
    - generally *differentially rotating*
    - relatively small random motions
- the Milky Way is highly flattened: is is a spiral galaxy

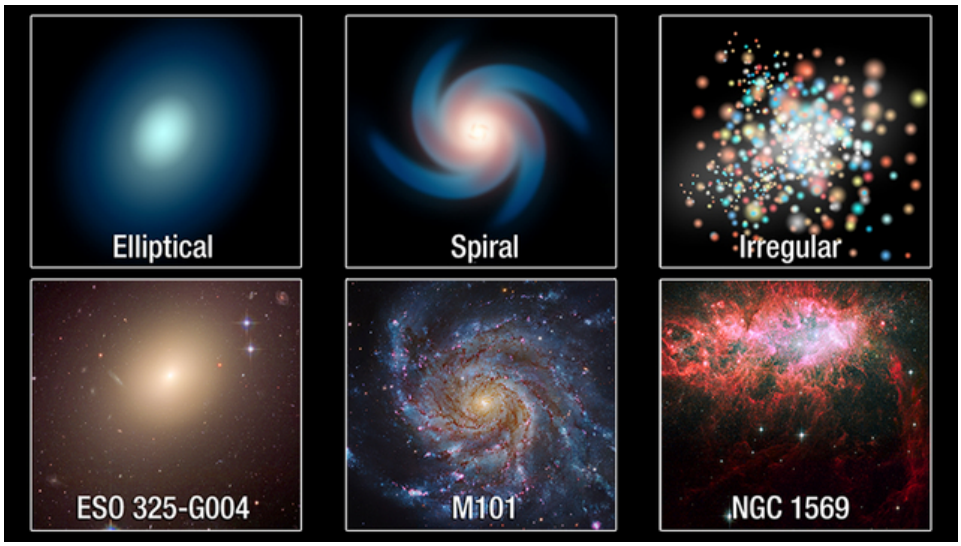


Figure 7: There are three main types of galaxies (though these are further classified morphologically). Of these, the bulk of the luminosity in the Universe comes from spirals and ellipticals. Spiral galaxies are supported against gravitational collapse by rotation. Elliptical galaxies are supported by 'kinetic pressure' (the motions of stars). The Milky Way (being highly flattened) is a spiral galaxy.

Q: How can we map the structure of the Milky Way using radio telescopes?
A: By mapping the HI intensity as a function of angle and frequency.

- In a differentially rotating disk there is a *mapping* between v_{los} and distance
  - mapping depends on direction (angle)
  - and v_{los} is directly measurable from the frequency via the Doppler shift
- reconstructing the spatial structure requires (at least) local knowledge of the rotation rate and the 'rotation curve'
  - i.e. how the rotation varies with galacto-centric distance
  - this was first established by Jan Oort.

3.2 Oort's constants

Stellar kinematics:

- Optical measurements of local stars can provide very precise angular positions and 'proper motions'
  - i.e. how the stars are moving on the sky (relative to a frame defined by distant galaxies or quasars)
- Proper motions show a steady linear trend plus an oscillating component
- the latter being the 'reflex' of the motion of the Earth around the sun
  - this 'parallax' determines the distance d
  - one can also get less precise distances from 'spectroscopic parallaxes'
    - where the distance is inferred from the flux density using an absolute velocity inferred from the spectrum of the star and stellar structure theory
- The distance and the steady component of the proper motion determine the 2 components of the tangential velocity V_t perpendicular to the line of sight
- The redshift of the spectral lines gives the radial component of the velocity V_r

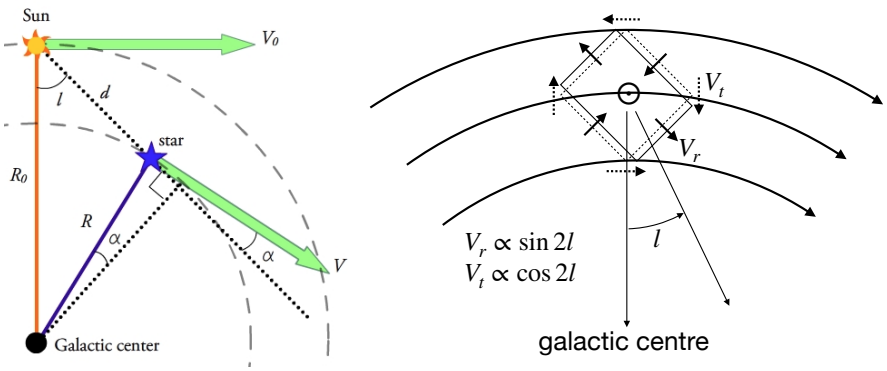


Figure 8: Left panel shows the geometry for Oort's calculation. In a rigidly rotating disk, stars rotate on the sky at fixed angular velocity $\Omega_0 = V_0/R_0$ with zero radial velocities V_r. Right panel shows – in a rotating frame – how, in a differentially rotating disk, stars initially around us on a diamond (dashed) will be 'sheared' into the rectangle (solid) and have non-vanishing $V_r \propto \sin 2l$.

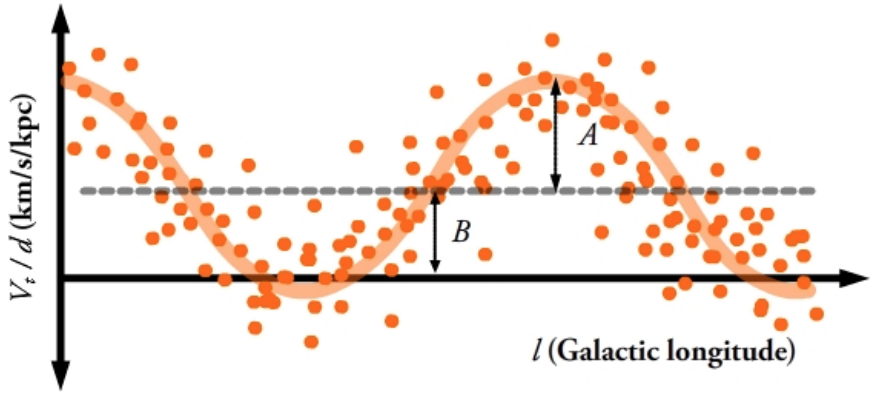


Figure 9: Illustration of how Oort's constants A and B can be inferred from noisy observations of proper motions of stars. These allow one to determine the local rotation $\Omega_0 = V_0/R_0$ and the differential rotation $d\Omega/dR$. These allow one to un-map radial velocities to distance around the sky.

Analysis of kinematic data in the MW (technicality):

- the motion of stars is not purely rotational
- there is a random component – gives a 'noise'
- and the Sun has a non-rotational motion too – gives a 'systematic error'
  - the 'noise' is removed by averaging
  - the systematic error is removed by determining the 'local standard of rest' (LSR)
    - i.e. how the sun is moving relative to local stars

    * velocity observations can be corrected into the 'LSR frame'

What can we learn about the galactic rotation from (relatively) local kinematics?

- What if the galaxy were rotating like a solid body?
  - with fixed angular rotation frequency Ω
- local stars would rotate on the sky (relative to distant objects)
  - so they would have relative transverse velocities V_t proportional to d
  - and zero line-of-sight velocities V_r
- More general, we expect the galaxy to be *differentially rotating* with $\Omega = \Omega(R)$
- the local rate of change $(d\Omega/dR)_0$ is measurable
  - either from V_r via the 'shearing' of the local flow of stars around us $\propto \sin 2l$ (see figure 8)
  - or from an additional proper motion proportional to $\cos 2l$ (see figure 9)
- Oort's constants: A and B
  - linear combinations of Ω_0 and $(d\Omega/dR)_0$
  - $A = -R_0(d\Omega/dR)_0/2$
  - $B = -R_0(d\Omega/dR)_0/2 - \Omega_0$
- Allowed Oort and collaborators to 'un-map' from HI emission maps as a function of frequency (the offset in frequency from the rest-frame frequency being the Doppler shift caused by V_r) to real space distances
  - inverse mapping fails in certain directions
  - but clearly shows that the Milky Way has spiral arms
  - departures from purely circular motion – e.g. 'peculiar velocities' associated with the spiral arms themselves – can also be seen
    * Bart Bok called these *'fingers of god'*
    * God here telling us that we are making an error!

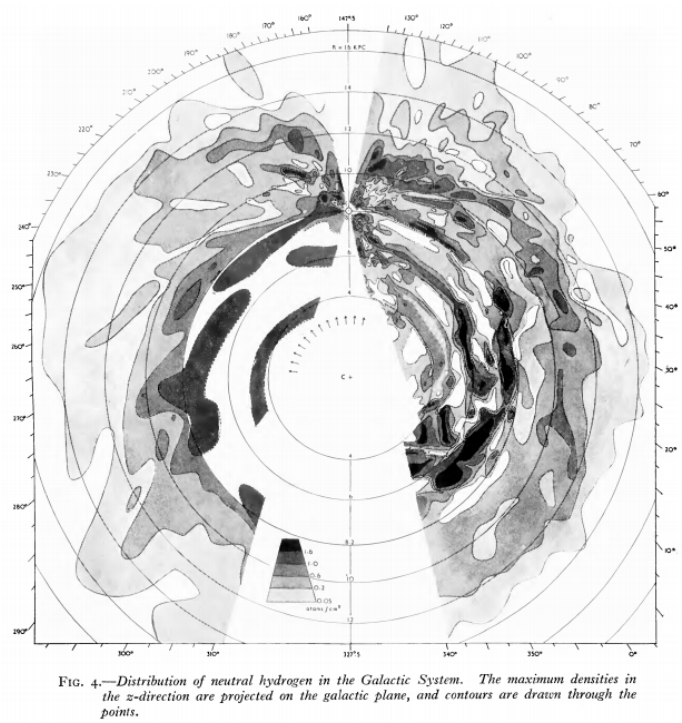


FIG. 4.—*Distribution of neutral hydrogen in the Galactic System. The maximum densities in the z-direction are projected on the galactic plane, and contours are drawn through the points.*

Figure 10: Oort's maps of the plane of the Milky Way reconstructed from measurements of the intensity $I_\nu(l)$ of the 21cm line (from Oort et al. 1958). The map shows the spiral structure of the MW. The paper also showed very clearly that the MW has a flat rotation curve. This was a surprise since it was expected that the rotation velocity would be falling as predicted by Kepler's law. Its failure to do so tells us that the outer parts of the galaxy are dominated by a dark matter halo.

4 Pulsar dispersion and Faraday rotation

4.1 The plasma frequency

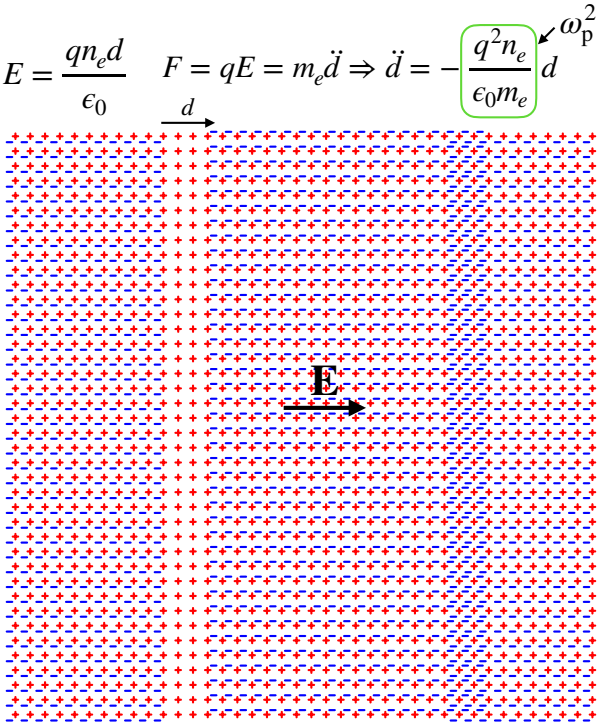


Figure 11: The Plasma Frequency

- Consider a uniform cold, non-magnetized plasma.
- Displace the electrons in a slab by a distance (perpendicular to the slab) d.
- This creates an electric field like that of a planar capacitor with charge surface density $\sigma = qn_e d$.
  - electric field: $E = \sigma/\epsilon_0$
  - gives restoring force proportional to d
  - so $\ddot{d} = -{\omega_\mathrm{p}}^2 d \quad \Rightarrow$ simple harmonic motion with (squared) frequency $\boxed{{\omega_\mathrm{p}}^2 = n_e q^2/\epsilon_0 m_e}$
- 1-D slab is a 'toy-model' – but result is general; any disturbance – of electrons relative to ions – will oscillate at frequency ω_p.

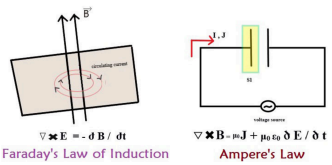
4.2 EM waves in a cold plasma

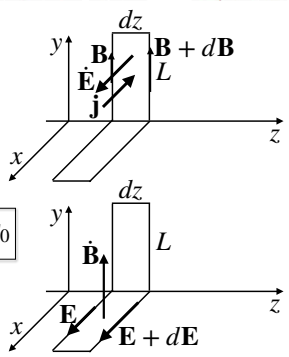
Recall EM waves in vacuum:

- for sinusoidally varying fields $E, B \sim \cos(kz - \omega t)$:
  - Faraday: $kE = \omega B$
  - Ampère/Maxwell: $kB = \omega E \quad \Leftarrow$ displacement current
  - satisfied if $\omega^2 = k^2$
- gives *wave speed* $c = \omega/k = 1$ (in 'sensible' units)
  - the speed with which wave crests, or nodes, move

Waves in cold plasma:

- $\mathbf{E} = \hat{\mathbf{x}} E_0 \cos(\omega t - kz)$
- $\mathbf{B} = \hat{\mathbf{y}} B_0 \cos(\omega t - kz)$ as before
- electrons feel acceleration
- $\ddot{x} = \frac{qE_0}{m}\cos \rightarrow \dot{x} = \frac{qE_0}{m\omega}\sin$
- so current is $\mathbf{j} = nq\dot{\mathbf{x}} = \frac{nq^2 E_0}{m\omega}\hat{\mathbf{x}}\sin = -\frac{nq^2}{m\omega^2}\dot{\mathbf{E}}$
- and $\dot{\mathbf{E}} + \mathbf{j}/\epsilon_0 = (1 - \omega_p^2/\omega^2)\dot{\mathbf{E}}$
- where plasma frequency is $\boxed{\omega_p^2 = nq^2/m\epsilon_0}$
- and dispersion relation is now
- $\boxed{\omega^2 = c^2k^2 + \omega_p^2}$ just like massive ϕ





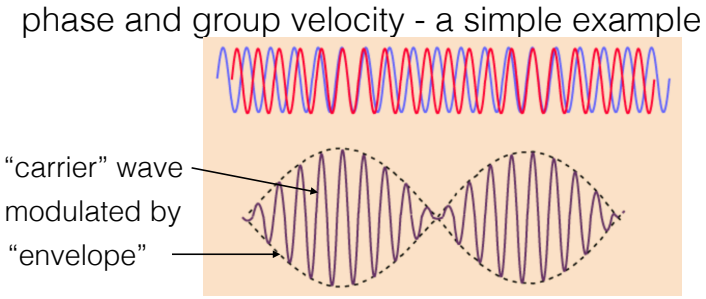
- In a plasma, the electric field will accelerate electrons and drive a current
  - the RHS of Ampère/Maxwell is reduced:
    - $\partial E/\partial t \rightarrow \partial E/\partial t + 4\pi j$
  - so $kB = \omega E(1 - \omega_p^2/\omega^2)$ and $kE = \omega B$
  - so $c = \omega/k = (1 - \omega_p^2/\omega^2)^{-1/2}$
- wave speed depends on frequency
  - no affect for $\omega \gg \omega_p$
  - speed for lower frequencies is changed
  - wave speed is 'super-luminal' $(c \geq 1)$!

4.3 Dispersive waves

EM waves in a vacuum – or waves on a violin string – are *non-dispersive*

- wave speed ω/k is independent of frequency
- admit 'd'Alembertian solutions: $f(x,t) = f_+(x-ct) + f_-(x+ct)$
- wave-forms propagate without changing shape

Wave-number $k = 2\pi/\lambda$. The relation between frequency and wave-number $\omega = \omega(k)$ is known as the *dispersion relation*. If it is 'non-trivial' (i.e. not just a constant multiple of k) then waves are said to be *dispersive*.



- trigonometry → $\cos a + \cos b = 2\cos\left(\frac{a-b}{2}\right)\cos\left(\frac{a+b}{2}\right)$
- so for travelling waves $a = (\overline{\omega} + \Delta\omega/2)t - (\overline{k} + \Delta k/2)x$, $b = (\overline{\omega} - \Delta\omega/2)t - (\overline{k} - \Delta k/2)x$
- $\cos a + \cos b = 2\cos(\Delta\omega t - \Delta k x)cos(\overline{\omega}t - \overline{k}x)$
- carrier $\cos(\overline{\omega}t - \overline{k}x)$ moves with "phase" speed $\overline{\omega}/\overline{k} \to \omega/k$
- envelope $\cos(\Delta\omega t - \Delta k x)$ moves with "group" speed $\Delta\omega/\Delta k \to d\omega/dk$

Figure 12: Group velocity

- Consider, for simplicity, the sum of two waves with slightly different wave-numbers k and $k' = k + \Delta k$ and with corresponding frequencies ω and $\omega' = \omega(k') = \omega + \Delta\omega$.
- simple trigonometry shows this is the product of two travelling waves
  - at half the sum and half the difference frequencies
- the key feature is that the *envelope* – the bunches or 'groups' of waves – travels at a *group speed* $v_g = \Delta\omega/\Delta k$
- and if $\Delta k \ll k$, $\Delta\omega/\Delta k \Rightarrow d\omega(k)/dk$

- argument can be generalised to case of *wave packets*
  - Fourier synthesis of a localised disturbance
    - if scale of variation of wave envelope is $\gg \lambda$ then sum is over waves with similar wavelength
    - and the 'packet' can propagate through many times its length without significantly changing its form
- the packet moves with speed $v_g \equiv d\omega/dk$: the *group velocity*
- while the wave-crests move at speed $v_p \equiv \omega/k$: the *phase velocity*
- over longer time-scales the different frequencies get dispersed

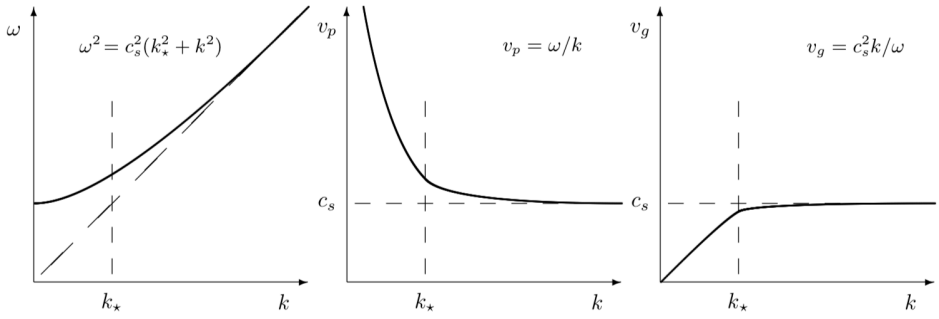


Figure 13: Dispersion relation, phase-velocity and group velocity for EM waves in a plasma. Here $k_\star$ is ω_p/c. The phase-velocity is super-luminal for all k, but this is no problem as no information travels at the phase-speed. Speed of information propagation is limited by the group-speed, which is sub-luminal for all k.

There are many interesting examples of dispersive waves in physics.

- In 2 dimensions the frequency $\omega_{\mathbf{k}}$ lies on a hyperboloid above/below the $\mathbf{k} = (k_x, k_y)$ plane
  - sometimes called the "mass shell"
- A wave $\phi(t, \mathbf{x}) = \frac{1}{2}(\phi_{\mathbf{k}} e^{i(\omega_{\mathbf{k}} t - \mathbf{k}\cdot\mathbf{x})} + \text{c . c.})$ has positive and negative frequency components
  - with opposite temporal and spatial frequencies

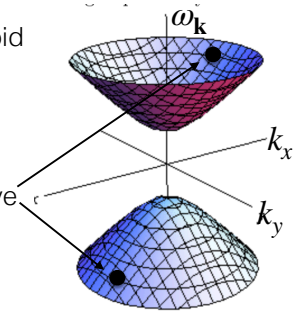


Figure 14:] The dispersion relation defines a hyper-surface in 4-dimensional frequency space: $\vec{k} = (\omega_{\mathbf{k}}, \mathbf{k})$. There are two separate hyper-surfaces for positive and negative $\omega_{\mathbf{k}}$. A real wave, beam or packet propagating (and transporting energy) in the direction $\hat{\mathbf{k}}$ is the sum of positive and negative frequency components $\propto e^{\pm i(\mathbf{k}\cdot\mathbf{x} - \omega_{\mathbf{k}} t)}$.

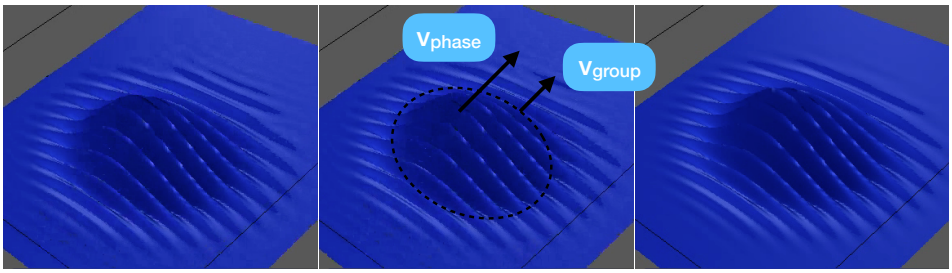


Figure 15: Some stills from an animation showing the propagation of a wave-packet made by summing waves with dispersion relation $\omega^2 = {\omega_{\rm p}}^2 + c^2 k^2$. The wave crests appear at the back of the packet and march through the packet, disappearing at the leading edge.

- De Broglie/Schrödinger waves
  - $E\psi = [(p^2/2m) + V]\psi \rightarrow \dot{\psi} = (\nabla^2 + a)\psi \rightarrow \omega(k) = a + k^2$
  - phase velocity $v_p = \omega/k = (a + k^2)/k$
    - can be superluminal: $v_p \rightarrow \infty$ as $k \rightarrow 0$
    - but carries no information
  - group velocity $v_g = d\omega/dk = 2k$
    - is well behaved as $k \rightarrow 0$
    - velocity of a packet in accord with idea that momentum is proportional to k.
  - an important application in astrophysics is in cosmology: a popular model for the DM is that it is the *axion field* or an *ultra-light axion-like* field.
    - in the potential wells of galaxies an clusters, these fields obey the non-relativistic Schrödinger equation
- Deep ocean water waves.
  - $\ddot{\psi} = \text{constant} \times \nabla\psi \quad \Rightarrow \quad \omega(k) = a\sqrt{k}$
  - so $v_p = \omega/k = a/\sqrt{k}$ and $v_g = d\omega/dk = a/2\sqrt{k} = v_p/2$
  - examples:
    - ripples from a pebble in a pond
    - or bow-waves from a boat
  - longer waves have higher phase velocity
  - wave-crests travel through a packet from stern to prow at twice the packet speed
  - the mathematics of these waves finds an important application in planetary atmospheres where they are called *'gravity waves'*
- 'Whistlers' in ionospheric physics.
- Slinky?

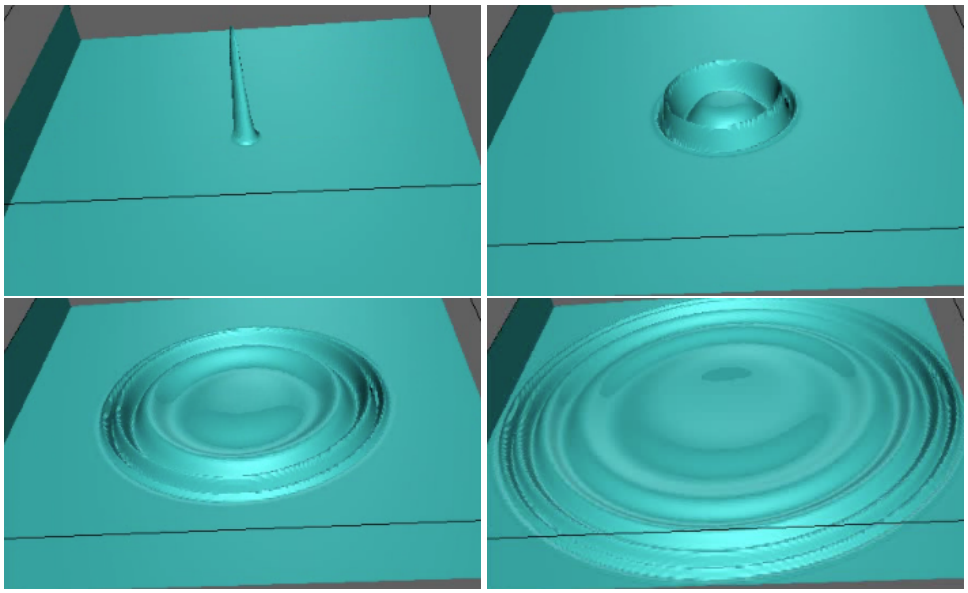


Figure 16: Some stills from an animation showing the Evolution of a 'pulse' in a cold plasma. The initial pulse contains components with a range of spatial frequencies. As it expands, the higher frequencies outrun the lower ones and the pulse becomes 'dispersed'. An observer at large distances will receive a 'pulse' with lower frequencies delayed. Radio astronomers observing pulsars use this to determine the integrated electron density along the line of sight.

Boat wake problem

- This is about phase and group velocities for deep ocean dispersive waves.
- Q1: what is the dispersion relation ?
- Force is (area $\sim L^2$) times density ρ times g times $h \leftarrow$ vertical displacement
- Mass is volume $\sim L^3$ times ρ
  - Assumes water down to depth $D \sim 1/k$ moves.
  - Why? No other scale in the problem
- EoM is $m\ddot{h} \simeq g\rho A h$ or $\ddot{h} \sim (g/L)h$ so frequency is $\omega \sim \sqrt{g/L}$
  - same as pendulum of length L
  - density doesn't appear
- so $\boxed{\omega = \alpha\sqrt{gk}}$ where α is a constant of order unity

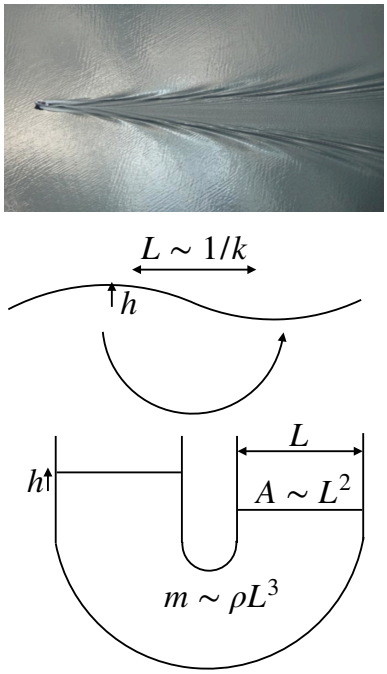


Figure 17: The 'bow-wake problem'. Here is an interesting little homework problem. You are sitting idly out on the ocean in a canoe when a big tanker comes past, generating a bow wave a bit like the one created by the speed-boat in the picture at left. There are N waves in the bow wave train. How many times will you bob up before the train passes you?

4.4 Refraction of EM waves in an inhomogeneous plasma

- The dispersion relation is $\omega^2 = {\omega_\mathrm{p}}^2 + k^2$
- so the wave-number is only real for frequencies ω in excess of the plasma frequency
- waves with $\omega < \omega_\mathrm{p}$ cannot propagate
  - if $\omega < \omega_\mathrm{p}$ we have imaginary $k = i\sqrt{{\omega_\mathrm{p}}^2 - \omega^2}$
  - so any (sensible) solution will be *evanescent* – i.e. falling off exponentially with distance
- In geometric optics we have *'Snell's law'* for refraction
  - it describes the abrupt change in ray directions at e.g. an air-glass interface in an optical system
  - and for a continuously varying refractive index $n(\mathbf{x})$ it says that the rate of change of the ray direction is the transverse gradient of $n(\mathbf{x})$
  - a result that follows from *'Fermat's principle of least time'* which says that the actual path of a ray is that which minimises the travel time

- the refractive index in optics is the inverse of the speed of light (in units of c)
- the situation for EM waves in an inhomogeneous plasma is analogous
  - beams, or wave-packets, of EM radiation propagating through a plasma with varying electron density will be deflected (refracted)
  - and will not be able to reach regions where they cannot propagate
- Q: What is the 'refractive index' (if we want Snell's law to still work)? It is c divided by some velocity, but which one – phase- or group-velocity?
- Q: If a plane wave propagating in empty space encounters a cloud of ionised plasma, does it get focussed or de-focussed?
- Q: Figure 18 shows a beam of radio waves getting reflected off the ionosphere. Sketch the wave-fronts in the beam as it undergoes reflection.

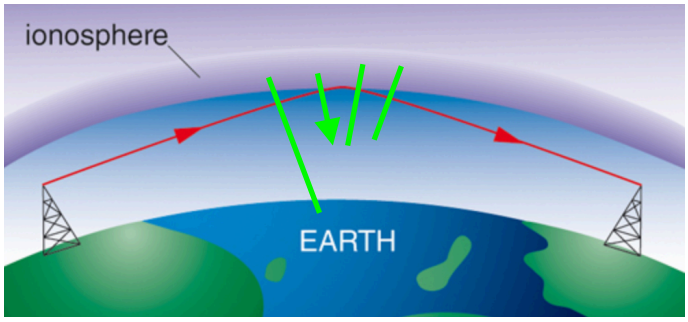
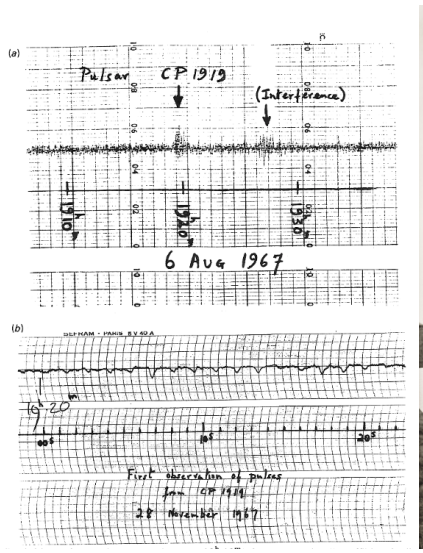


Figure 18: An application of the EM dispersion relation in a plasma is that of reflection of 'short-wave' radio waves off the ionosphere. The number density of electrons increases with altitude as one enters the ionosphere, as does the plasma frequency $\omega_{\rm p} \propto \sqrt{n_e}$. It peaks at $\nu = \omega_{\rm p}/2\pi \simeq 30$MHz. Only waves above this frequency can propagate through the atmosphere. This explains why the terrestrial window for radio-astronomy terminates at 30MHz. And how we can listen to distant short-wave radio stations.

4.5 Pulsar dispersion: the 'dispersion measure'



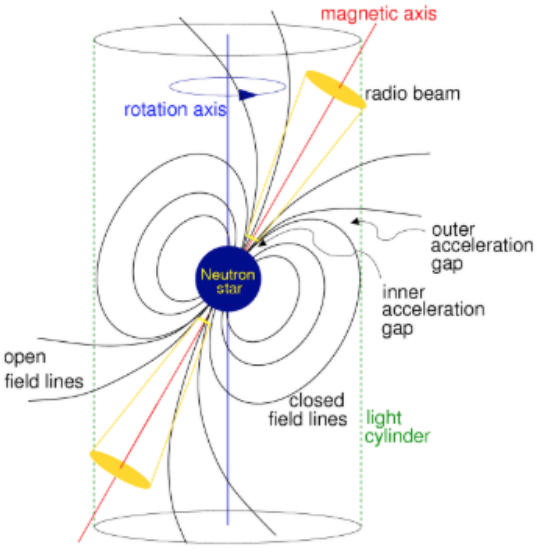


Figure 19: The first pulsar CP1919 was discovered in 1967 at Cambridge by Jocelyn Bell Burnell and her supervisor Anthony Hewish. Pulsars are believed to be magetised neutron stars in which electrons are accelerated and emit radio waves in a relatively narrow beam that sweeps around rather like the beam of a light-house. An observer in the path of the beam sees regularly space pulses. Pulsars, and particularly millisecond pulsars, are extremely stable 'clocks'. Some have been observed for tens of years and have pulse arrival times 'residuals' on the order of 100 ns. That's as stable as the best terrestrial atomic clocks. They also provide an extremely powerful probe of the ISM.

- With the dispersion relation
  - $\omega = \sqrt{c^2k^2 + {\omega_\mathrm{p}}^2}$
- the group velocity is
  - $v_\mathrm{g}/c = c^{-1}d\omega/dk = \sqrt{1 - {\omega_\mathrm{p}}^2/\omega^2} = \sqrt{1 - (n_e q^2/4\pi^2\epsilon_0 m_e)\lambda^2}$
  - so longer wavelength radiation travels slower
- this means that a pulse of radiation from a distant source will become a 'chirp'
  - where the highest frequencies arrive first
  - followed by a tail of lower frequencies
- Pulsars – as their name suggests – emit pulses
- and these do indeed appear dispersed – lower frequencies arrive later
- measurement of arrival times at different frequencies gives a quantitative 'measure' of an integral of the electron density
  - the arrival time is
    - $t(\lambda) = t_\mathrm{em} + \int dz/v_\mathrm{g}(z, \lambda)$
    - where z is distance along the line of sight
  - for frequencies much greater than the plasma we can expand
    - $c/v_\mathrm{g} = (1 - (n_e q^2/4\pi^2\epsilon_0 m_e)\lambda^2)^{-1/2} \simeq (1 + (n_e q^2/8\pi^2\epsilon_0 m_e)\lambda^2)$
  - so the pulse arrival time delay
    - relative to that at high frequency
    - or relative to the arrival time at some arbitrary fiducial frequency, is
    - $\delta t(\lambda) = (q^2/8\pi^2\epsilon_0 c m_e)\lambda^2 \int dz n_e(z)$
  - so simply differentiating $\delta t(\lambda)$ with respect to λ^2 gives the *dispersion measure:*
    - $\mathrm{DM} \equiv (q^2/8\pi^2\epsilon_0 c m_e) \int dz n_e(z)$
  - pulsars provide a quantitative measure the integrated plasma density in the Milky Way

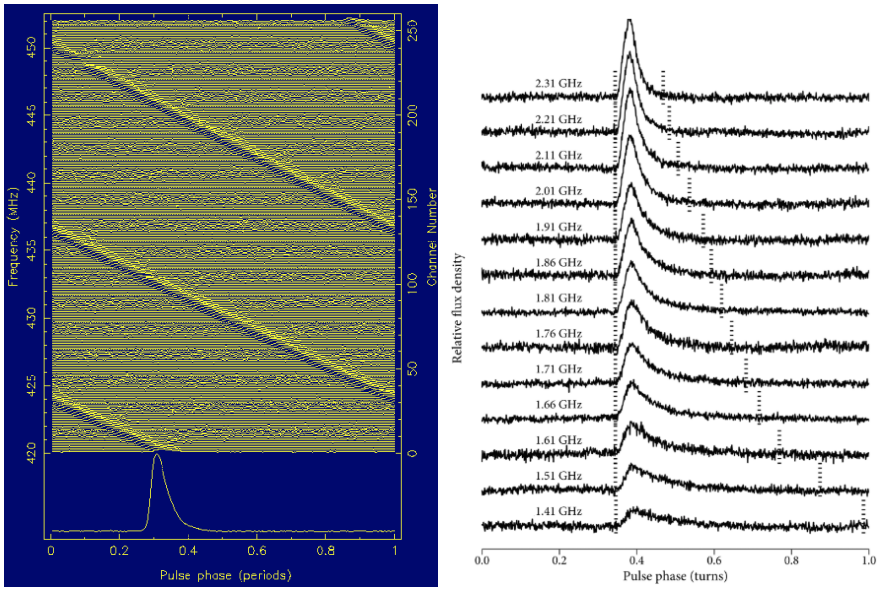


Figure 20: Left panel shows the pulse arrival time vs. frequency for a pulsar. The output of the radio telescope has been split into 256 different frequency channels to give the flux density as a function of frequency and time. The time (horizontal) axis is actually the phase; the arrival times have been converted to a phase by dividing by the period of the pulses and then wrapped. This is done solely to increase the signal to noise. The curve at the bottom is the 'un-dispersed' pulse profile. Right panel shows that, in some cases, the pulse profile gets broader at lower frequencies.

4.6 Faraday rotation: the 'rotation measure'

A related diagnostic of the ISM (or IGM) comes from Faraday rotation.

- radio sources are usually highly linearly polarized (synchrotron or cyclotron emission, for example).
- If there is a component of the magnetic field parallel to the line of sight, the plane of linear polarisation rotates because of the Faraday effect.
- it may be difficult to say what the intrinsic polarisation angle is is
- but, like dispersion, the effect is frequency dependent – it is greater at low frequencies
- so measurements at different frequencies – which reveal how the polarization angle *changes* with frequency – give the *rotation measure*
  - an integral of electron density times magnetic field
- combining with dispersion measurement gives estimate of B
- but complicated as reversal of the field causes cancellation

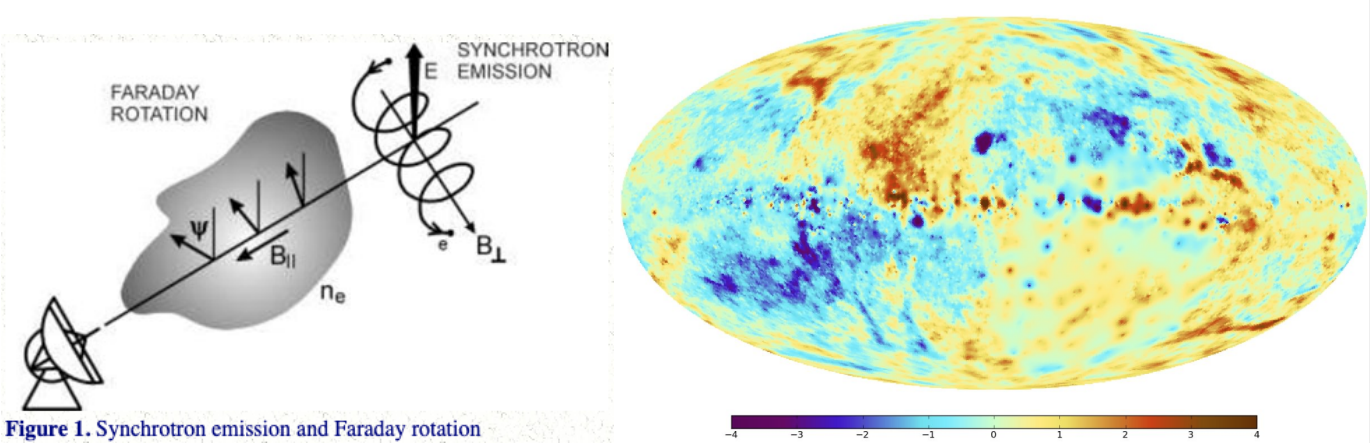


Figure 1. Synchrotron emission and Faraday rotation

Figure 21: Left panel shows how the polarisation angle of radiation from a distant radio source gets rotated as it passes through the ISM. The amount of rotation gives the line integral of the electron density n_e times the line-of-sight component of the magnetic field $B_{\|} \equiv \hat{\mathbf{n}} \cdot \mathbf{B}$. The right hand side is a map of the rotation measure in galactic (l, b) coordinates.

5 Kolmogorov theory of turbulence

Turbulence is observed in fluids in a wide range of situations and is near ubiquitous in the ISM. As this is an essentially 3D phenomenon, and the equations of fluid dynamics are non-linear, one might think that this could only be addressed by numerical simulations.

However, in the early 40s, Andrey Kolmogorov developed a statistical approach to the problem. He was able to show that, provided there is a large span between the large-scale (the 'outer-scale') on which kinetic energy is being injected and the 'inner-scale' at which molecular viscosity is damping the microscopic eddies, the will be a power-law-like behaviour for things like the two-point correlation function for the velocity field. The simplest models are somewhat crude, and make various assumptions that may or may not be valid. Nonetheless, they seem to explain quite well a broad range of phenomena, such as scintillation – which provides a very powerful probe of the structure of the interstellar plasma.

Below we will first discuss the expected structure of the turbulent ISM in Kolmogorov's universal scaling model. We then describe the structure of the wave-front corrugation that this produces in the radiation from distant point-like sources and finally we describe the different types of scintillation mentioned above.

5.1 Structure of the turbulent ISM

- There are good reasons to expect that the interstellar plasma is turbulent
  - it is subject to pressure and magnetic forces
  - its viscosity is negligible for large-scale motions
  - so it has a very large *Reynold's number*
    - Euler equation for a viscous fluid:
    - $\rho dv/dt = -\nabla P + \sigma \nabla^2 v$
    - where σ is the coefficient of viscosity
    - Reynold's number: $R = \rho L v/\sigma$ indicates relative (un)importance of viscosity
- so we might expect it to behave a lot like what happens if we stir an inviscid fluid
  - we start with smooth motion on a large scale
  - this is called the *'outer scale'*: the scale on which we are injecting energy
  - so initially the power-spectrum of motions would be concentrated at low-k (spatial frequency)
  - but just as sketched by da Vinci and seen in everyday life there is a *'turbulent cascade'* where big eddies create smaller eddies and these in turn create smaller eddies and so on until, ultimately the eddies are of such a small scale that particle diffusion is able to damp out the motions (at the *'diffusion scale'*)
  - mathematically, this can be understood qualitatively as arising as a result of *non-linearity* in the Euler (or Navier-Stokes) equation, where the term representing the acceleration of an element of fluid is $d\mathbf{v}/dt = (\partial_t + \mathbf{v}\cdot\boldsymbol{\nabla})\mathbf{v}$
    - for sound waves in a gas, for instance, the acceleration is minus the gradient of the pressure – a *force density* – divided by the density
    - if the amplitude of the density fluctuations is small, then we can ignore $(\mathbf{v}\cdot\boldsymbol{\nabla})\mathbf{v}$ (being second order in the amplitude of the wave) as compared to $\partial_t \mathbf{v}$
    - and, combining with the continuity equation $\dot{\rho} + \boldsymbol{\nabla}\cdot(\rho\mathbf{v}) = 0$, and with some equation of state, we get a *linear* wave equation for $\delta\rho(\mathbf{x},t) \equiv \rho(\mathbf{x},t) - \overline{\rho}$
    - which we can solve by making a Fourier decomposition $\delta\rho(\mathbf{x},t) = \sum_{\mathbf{k}} \delta\rho_{\mathbf{k}} e^{i(\mathbf{k}\cdot\mathbf{x} - \omega_{\mathbf{k}} t)}$
    - with the different $\mathbf{k}$-modes evolving independently
  - when the amplitude becomes large, however, we cannot ignore the non-linear term
    - this *couples* the density and velocity fluctuations at different scales
    - if we start with large-scale motions (at the outer scale $k_{\rm outer}$) the non-linearity will excite fluctuations at smaller scales (higher k).
    - this is what is happening in a turbulent fluid
- as first analysed by Kolmogorov this produces universal statistical scaling laws structure for the velocity and density power spectra

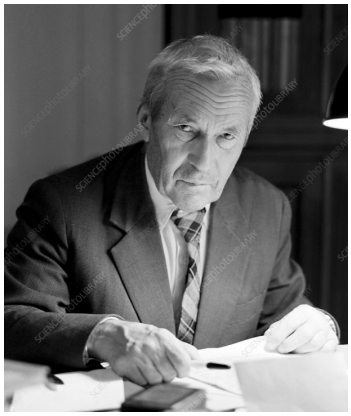


Figure 22: Left: Sketch by Leonardo da Vinci. Right: Soviet mathematician Andrei Kolmogorov who first derived the universal scaling laws for turbulent fluids. One of these is that the velocity v of eddies of size L scales as $L^{1/3}$.

5.1.1 Scaling of velocity with eddy size

- Figure 23 illustrates what happens if we 'inject' – in a transient manner – kinetic energy at the outer scale. After about one 'eddy turnover time' $\tau \sim L/v$, we will have smaller eddies. And these will generate smaller eddies, until, ultimately, the energy will be dissipated at the diffusion scale.
- In a situation such as that sketched by da Vinci (figure 22), where we have *continual* injection of energy at the outer scale, and diffusion dissipating the energy at the inner scale – turning it into heat – we expect to obtain what is called *'fully-developed turbulence':*
  - At each scale L we have eddies with characteristic velocity v_L with energy density $u_L \sim \rho v_L^2$. Non-linearity will transfer this energy into eddies of smaller scale (say a factor 2 smaller – or some similar number of order unity) on a time-scale $\tau_L \sim L/v_L$. In the steady state, this will be replenished by energy being transported from the next larger scale.
- the picture here is of a *'hierarchy'* of eddies – equally spaced in $\log L$
- if a steady state develops it must be that the rate at which eddies are transferring energy to their offspring must be independent of scale:
  - $P_L \equiv u_L/\tau_L \sim (\rho v_L^2)/(L/v_L) \propto L^0$
- which implies
  - $\boxed{v_L \propto L^{1/3}}$

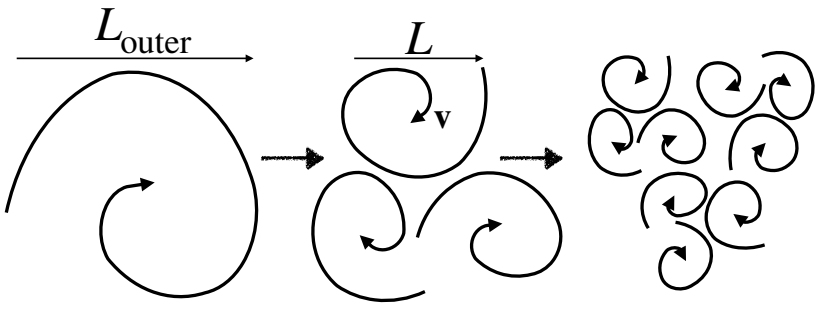


Figure 23: The idea behind Kolmogorov scaling is that eddies of size L with velocity v_L will generate smaller eddies on a time-scale $\tau_L \sim L/v_L$. The kinetic energy density is $u_L \sim \rho v_L^2$. So the eddies of size $\sim L$ are transferring energy per unit volume to smaller scales at a rate $P_L \sim \rho v_L^2/\tau_L \propto v_L^3/L$. So if energy is not to build up at a particular scale, we need $P_L \propto L^0 \Rightarrow v_L \propto L^{1/3}$.

5.1.2 Structure function for 'passive additives'

- An analogous argument allows us to predict the power spectrum $P(k)$ of fluctuations in density of a *'passive additive'*
  - an example of which might be the fluctuations $\delta_c(\mathbf{x}, t)$ in the density $c(\mathbf{x}, t)$ of cream in a cup of coffee that has been stirred
  - initially, the power will be entirely at low-k
    - the power being defined to be $P_c(k) = \langle|\tilde{\delta}_c(\mathbf{k})|^2\rangle$
    - where $\tilde{\delta}_c(\mathbf{k}) \equiv \int d^3x\, \delta_c(\mathbf{x}) e^{i\mathbf{k}\cdot\mathbf{x}}$ is the Fourier transform of $\delta_c(\mathbf{x})$
    - and which obeys *Parseval's theorem* $(2\pi)^{-3} \int d^3k P_c(k) = V^{-1} \int d^3x\, \delta_c^2(\mathbf{x})$
    - which means that $k^3 P_c(k)$ measures the variance per log-interval of wave-number k
  - but as sub-eddies develop, there will be a cascade of power to higher k
- and again, if we are dealing with *fully developed turbulence*
  - i.e. with continual injection of inhomogeneity in the additive at the outer scale
    - one might think of dye being injected
  - and with continual destruction of inhomogeneity by diffusive mixing at the inner scale

- the variance is being transported down the hierarchy of eddies and the variance at scale $L \sim k^{-1}$
  - $\langle \delta_c^2 \rangle_L \sim (k^3 P_c(k))_{k \sim L^{-1}}$
- divided by the eddy turnover time $\tau_L \sim L/v_L \propto L^{2/3}$ must be independent of L
  - $\boxed{(k^3 P_c(k))_{k \sim L^{-1}} \propto L^{2/3}}$
- in studies of turbulence, as well as the power spectrum $P_c(k)$, it proves to be useful to describe the statistical properties of the fluid in terms of the *'structure function':*
  - $\boxed{S_c(\mathbf{L}) \equiv \langle (\delta_c(\mathbf{x}) - \delta_c(\mathbf{x}+\mathbf{L}))^2 \rangle_{\mathbf{x}}}$
  - where $\langle \ldots \rangle_{\mathbf{x}}$ denotes the spatial average
- which, for *isotropic* turbulence, is only a function of $L = |\mathbf{L}|$
  - the structure function is closely related to the *autocorrelation-function:*
    - $\xi_c(\mathbf{x}) \equiv \langle \delta_c(\mathbf{x}') \delta_c(\mathbf{x}'+\mathbf{x}) \rangle_{\mathbf{x}}$
  - since $S_c(\mathbf{x}) = 2(\xi_c(0) - \xi_c(\mathbf{x}))$
  - and the Wiener-Khinchin theorem says that $P_c(k)$ is the Fourier transform of $\xi_c(x)$
  - so $P_c(k)$, $\xi_x(x)$ and $S_c(x)$ all provide equivalent information about $\delta_c(\mathbf{x})$
  - we say they all describe the *second moment* of the fluctuations
- in terms of $S_c(L)$, the Kolmogorov scaling law for a passive additive is
  - $\boxed{S_c(L) \propto L^{2/3}}$

5.1.3 Specific entropy and density as passive additives

- A particularly important passive additive is the *specific entropy* $s(\mathbf{x})$ (the entropy per particle) of a gas or plasma
  - assuming that we are dealing with eddies sufficiently large that heat conduction is negligible the structure function $S_s(L)$ should obey Kolmogorov's scaling law $S_s(L) \propto L^{2/3}$
- and, if we are dealing with *sub-sonic turbulence*, so the gas can be considered to be effectively *isobaric*, then the fluctuations in the density n of gas particles (or n_e for electrons in a plasma) will also obey $S_n(L) \propto L^{2/3}$
  - the specific entropy is $\boxed{s = k_{\rm B} \log\left(T^{3/2}/n\right)}$
    - since if we heat an element of gas at constant n we have $dS = VdQ/T = Vd((3/2)nk_{\rm B}T)/T = (3/2)nVk_{\rm B}d\log T$ so $ds = dS/nV = (3/2)k_{\rm B}d\log T$ so at contant n (or constant V)
    - while if we change the volume adiabatically ($dQ = 0$) the pressure is $P = nk_{\rm B}T \propto V^{-5/3} \propto n^{5/3}$ implying $T \propto n^{2/3}$ at constant s, so s must be a function only of the combination $T^{3/2}/n$
    - another way to see this is to use the statistical mechanical entropy $S = -\int d^3x \int d^3p f(\mathbf{p}) \log f(\mathbf{p})$ and $N = \int d^3x \int d^3p f(\mathbf{p})$ so $s = S/N = -\langle \log f \rangle$. But the phase-space density f is on the order of $f \sim n/\langle v^2 \rangle^{3/2}$
      - i.e. the number per spatial volume per volume in velocity space
    - and $\langle v^2 \rangle \propto T$, so again we get $s = \log f^{-1} = \log\left(T^{3/2}/n\right) + \text{constant}$
    - aside: it is common in e.g. atmospheric physics to refer to the specific entropy of a body of gas in terms of the *potential temperature*: this being the temperature the gas would have if it were brought, adiabatically, to some standard reference density
  - so $ds/k_{\rm B} = (3/2)dT/T - dn/n$
  - while, at constant pressure, $dP = d(nk_{\rm B}T) = nk_{\rm B}T(dn/n + dT/T) = 0 \Rightarrow dT/T = -dn/n$
  - hence $ds/k_{\rm B} = -(5/2)dn/n$ implying

    - $\boxed{dn = -(2n/5k_{\rm B})ds}$
  - so the fluctuations in density are linearly proportional to the fluctuation in the specific entropy
  - and the same will be true of the fluctuations in the refractive index (which is what concerns us here)
  - so both are expected to have Kolmogorovian 3D structure functions

5.2 Structure of wave-front corrugation caused by the inhomogeneous ISM

- We have described the Kolmogorov scaling for 3D structure function of e.g. the refractive index n:
  - $\boxed{S_n(L) \propto L^{2/3}}$
- what we are more interested in is the two-dimensional structure function for the height $h(\mathbf{x})$ of corrugations of the wave-fronts of light from distant sources that have passed through a 'slab' of turbulent plasma with such a structure function
- As described in the caption to figure 24 these corrugations have a 2D structure function
  - $\boxed{S_h(L) \propto L^{5/3}}$
  - another, more formal, way to arrive at this result is to use *Limber's equation*
  - and another way to convince yourself that this is right is to think of a *'white-noise'* (or *'shot-noise'* process – i.e. the kind of fluctuations in density you get by randomly distributing particles in cells)
    - for such a process the RMS density fluctuation $\delta_{3D}(L)$ in a 3D cell scales as $1/\sqrt{\overline{N}}$, where $\overline{N}$ is the mean count per cell, or as $L^{-3/2}$
    - while the projected RMS density fluctuation scales $\delta_{2D}(L)$ scales inversely with the *area* of the projected column of 3D cells, or as $\delta_{2D}(L) \propto 1/L$
    - the structure functions (in 2- and 3-D) being the variances we have $S_{2D}(L) \propto 1/L^2$ and $S_{3D}(L) \propto 1/L^3$ so we have, for the scaling with cell-size, $S_{2D}(L) \propto LS_{3D}(L)$
    - but if we didn't have white-noise scaling in 3D (as is the case for passive additives with Kolmogorov statistics), this will just 'tilt' both structure functions
    - another consequence of this line of argument is that if we have a 3D power-law power spectrum $P_{3D}(k) \propto k^n$
      - where n is often called the *'spectral index'*
      - and for the specific case of white-noise $n = 0$
    - then the power spectrum $P_{2D}(k)$ of the projected density will have the *same* index as the 3D power spectrum
      - which for the case of passive additives with Kolmogorov scaling is $\boxed{n = -11/3}$
      - since $S_{3D}(L) \sim (k^3 P_{3D}(k))_{k\sim 1/L} \propto L^{2/3} \sim k^{-2/3}$
      - or equally, since $S_{2D}(L) \sim (k^2 P_{2D}(k))_{k\sim 1/L} \propto L^{5/3} \sim k^{-5/3}$
- another feature of these wavefront corrugations is that, being the sum of statistically independent random spatial fluctuations, we would expect, by virtue of the *central limit theorem* that they will obey *Gaussian statistics*
  - one consequence of which is that one can generate realisations by adding waves (Fourier synthesis) with *randomly chosen phases*
  - another consequence is that for a *statistically homogeneous random field* the power-spectrum (or auto-correlation function or structure function) gives a *complete* description of the field

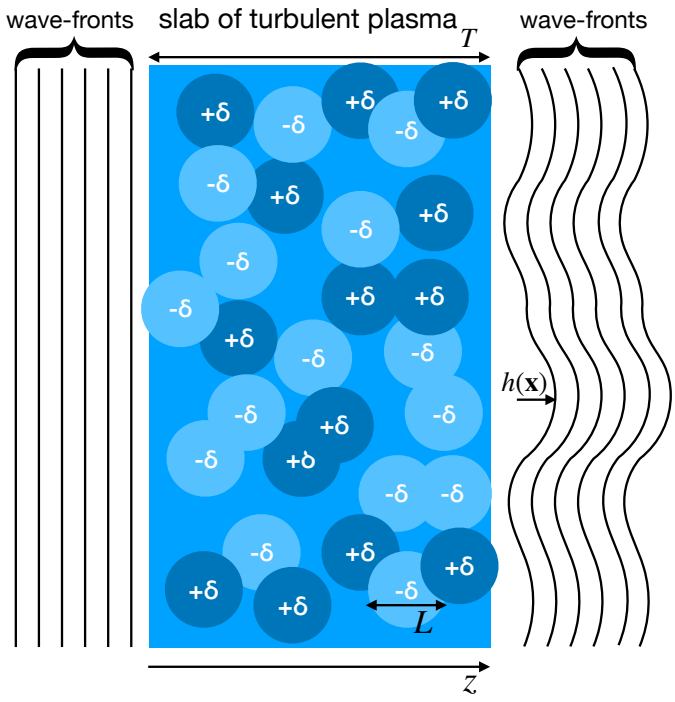


Figure 24: Kolmogorov scaling for the corrugation of wavefronts. The wavefront displacement is $h(\mathbf{x}) = -\int dz \delta n(\mathbf{x}, z)$. Consider a simple model of a slab of thickness T filled with randomly placed 'blobs' of size L with randomly chosen positive or negative fluctuations in the refractive index $\delta n = \pm\delta$. Each of these will give a contribution to $h(\mathbf{x})$ equal to $h_1 = \pm L\delta$. There will be, on average $N = T/L$ of these along any line of sight through the slab. So the mean-squared corrugation will be $\langle h^2 \rangle = N\langle h_1^2 \rangle = N(L\delta)^2 = TL\langle \delta^2 \rangle$. If we consider the superposition of a hierarchy of blobs of different sizes, and let the blobs of size L have refractive index obeying the Kolmogorov scaling law, so $\langle \delta_L^2 \rangle \propto L^{2/3}$, then the structure function of $h(\mathbf{x})$ on scale L – being driven by those blobs of size L – will scale as $S_h(L) \propto L\langle \delta_L^2 \rangle \propto L^{5/3}$. So the RMS difference in h at two points separated by distance L is $\delta h(L) = \sqrt{S_h(L)} \propto L^{5/6}$. That means that the 'tilt' of the wavefront $\delta h(L)/L$ is almost scale invariant.

6 Scintillation

Pulsar dispersion measures the *mean* electron density along a line of sight. A somewhat different probe of the ISM – one which measures the *inhomogeneity* or 'clumpiness' of the plasma – comes from *scintillation* of radio sources. This is variation of the flux density – quite analogous to the 'twinkling' of stars seen through Earth's atmosphere – caused by inhomogeneity of the plasma through which the radiation propagates. Interestingly, the first pulsar CP1919 was discovered in a survey carried out to look for scintillation of sources. The picture we have in mind is a distant point, or point-like, source and some 'scattering screen' of inhomogeneous ISM material that imposes corrugations on the wave-fronts from the source and which give rise to spatial fluctuations in the flux density down-stream of the scattering screen. The observable time variation of the flux-density arises because of our motion (or, more generally, the relative motions of the source, screen and observer).

As we will describe, the strength and character of the scintillation depends essentially on the relative size of two length scales:

- one is called the *'diffraction-scale'* $r_{\rm diff}$ which is the scale over which the inhomogeneous ISM produces phase errors on the order of one radian.
  - in atmospheric 'seeing' this is called the *'Fried length'*, after David Fried.
- the other is the *'Fresnel scale'* $r_{\rm F} \sim \sqrt{\lambda D}$, determined by the wavelength and the distance to the source
  - this is the size of a region on the scattering plane for which radiation from the source interferes constructively

The regimes $r_{\rm diff} \gg r_{\rm F}$ and $r_{\rm diff} \ll r_{\rm F}$ are known as the *'weak scattering'* and *'strong scattering'* regimes respectively.

- in the weak scattering regime, the amplitude of the flux density fluctuations is a diffraction phenomenon and the flux-density fluctuations are small
- in the strong scattering regime the flux-density fluctuations are strong and are caused by multi-path propagation with interference between the different paths
- there can also be longer time-scale *'refractive'* flux density fluctuations which, as the name suggests, can be understood in terms of geometric- or ray-optics which may be strong if caustics are formed.

Observed scintillation is broadly in line with the idea that the ISM is *turbulent*, with the characteristic power-law spectrum of density fluctuations as worked out, for the case of 'fully-developed isotropic turbulence' by Kolmogorov. Though one of the main goals in scintillation studies is to test for departures from the idealised model (as caused by anisotropy associated with magnetic fields, for instance).

In addition to providing information about the ISM, scintillation provides useful information about the relative velocities of the source, observer and scattering inhomogeneities as well as about the physical size of the emitting region.

6.1 The Fresnel-Kirchhoff integral

An excellent and accessible reference for the following material is Ramesh Narayan's 1992 paper in Phil. Trans. R. Soc. Lond. A.

According to Fresnel theory, the flux density at position $\mathbf{x}$ on the observer plane is given by $F(\mathbf{x}) = |\psi(\mathbf{x})|^2$ where the complex field amplitude $\psi(\mathbf{x})$ is given by the *'Fresnel-Kirchhoff integral'*:

$$\boxed{\psi(\mathbf{x};D) = \frac{i}{2\pi r_{\rm F}^2}\int d^2q \exp\left(i\phi(\mathbf{q}) + i|\mathbf{q}-\mathbf{x}|^2/2r_{\rm F}^2\right)} \tag{1}$$

where $\phi(\mathbf{q}) = 2\pi h(\mathbf{q})/\lambda$ is the phase-error ($h(\mathbf{q})$ being the corrugation of the wavefront) and where $r_{\rm F} \equiv \sqrt{\lambda D/2\pi}$, D being the distance from the observer plane to the scattering screen (see figure 25).

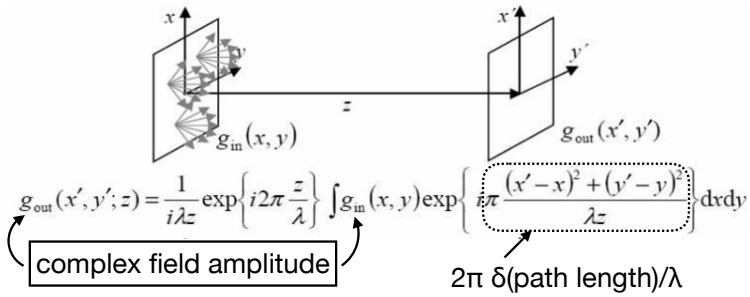


Figure 25: The Fresnel-Kirchhoff integral: The electric (or magnetic) field from a distant source at some position $\mathbf{x}$ on the plane $z = 0$ is the real part of a complex amplitude (here denoted by $g(\mathbf{x})$) times $e^{i\omega t}$. The amplitude $g(\mathbf{x}')$ on another plane – say the observer plane – is the sum of the amplitudes of all the elements of area on the first plane times a *'geometric phase factor'* $e^{i\pi|\mathbf{x}'-\mathbf{x}|^2/\lambda z}$.

Equivalently, the complex field amplitude at the observer plane $\psi(\mathbf{q};D)$ is the convolution $\psi(\mathbf{q};0) \otimes f(\mathbf{q};D)$ of the complex amplitude immediately down-stream of the scattering screen: $\psi(\mathbf{q};0) = e^{i\phi(\mathbf{q})}$ with the Fresnel function $f(\mathbf{q};D) = \exp\left(-|\mathbf{q}|^2/2r_{\rm F}^2\right)$.

A realisation of a phase-screen $\phi(\mathbf{q})$ as generated by Kolmogorov turbulence is shown as the lower surface in figure 26 and the upper surface shows the real part of the Fresnel function. To calculate the field amplitude for an observer at $\mathbf{x} = 0$, we multiply the complex exponential of the phase screen by the Fresnel function and integrate. And to get the normalised flux density we square the absolute modulus. To get the result at some different observer position we simply translate the Fresnel function and repeat.

6.2 Scintillation in the weak scattering regime

- The phase structure function for Kolmogorov turbulence is
  - $\langle(\phi(\mathbf{x}) - \phi(\mathbf{x}+\mathbf{r}))^2\rangle = S_\phi(r) = (r/r_{\rm diff})^{5/3}$
  - so $r_{\rm diff}$ is the *'phase coherence length'*: the rms phase fluctuation between two points with separation $r_{\rm diff}$ is 1 radian
- The *'weak scattering regime'* is $r_{\rm diff} \gg r_{\rm F}$
- this implies that the fluctuations in phase over the Fresnel scale are small and we have
  - $\psi(\mathbf{q}; z=0) = \exp(i\phi(\mathbf{q})) \simeq 1 + i\phi(\mathbf{q}) + \ldots$
  - or $\psi(\mathbf{q}; z=0) = 1 + \delta\psi(\mathbf{q}; z=0)$ with initial amplitude perturbation at linear order
  - $\boxed{\delta\psi(\mathbf{q}; z=0) = i\phi(\mathbf{q})}$
- So the perturbation to the field amplitude on a plane at distance z down-stream is

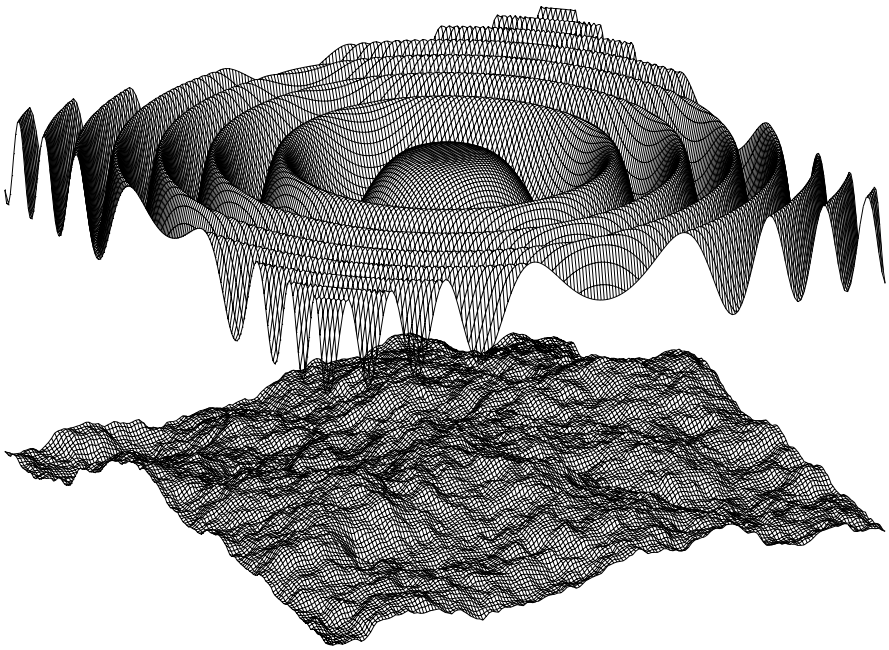


Figure 26: A realisation of a Kolmogorov 'phase screen' $\phi(\mathbf{q})$ (bottom) and (real part of) Fresnel function $f(\mathbf{q})$ top. The EM field immediately after passing through the scattering layer is that of an unperturbed plane wave $e^{i\omega t}$ times the complex amplitude $\psi(\mathbf{q}; z = 0) = e^{i\phi(\mathbf{q})}$. At a distance $z = D$ down-stream, the complex amplitude $\psi(\mathbf{q}; z = D)$ is given by the convolution of $\psi(\mathbf{q}; z = 0)$ with the Fresnel function, which aside from a normalisation factor, is $f(\mathbf{q}) = e^{i|\mathbf{q}|^2/2r_F^2}$. The convex bowl in the centre of the $f(\mathbf{q})$ is known as the first Fresnel zone.

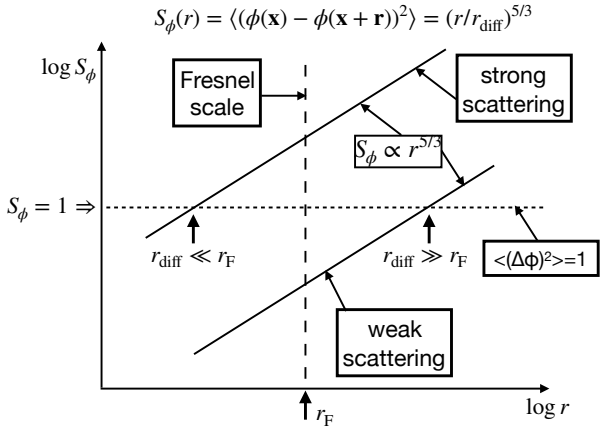


Figure 27: The *character* of the scintillation of a source depends on the relation between two scales: the *Fresnel scale* $r_F \sim \sqrt{\lambda D}$, whose significance is that the geometric path length difference for paths separated by less than r_F is less than a wavelength, so such paths interfere constructively, and the *diffraction scale* $r_{\rm diff}$, which is the scale at which phase fluctuations are of order 1 radian. In the strong (weak) scattering regimes, the phase fluctuations at the Fresnel scale and on most scales of interest are large (small). In the weak regime the scintillation is strongest at the Fresnel scale, but is weak. In the strong regime, large flux density modulation is found at the diffraction scale.

  - $\delta\psi(\mathbf{x}; z) = -\int \frac{d^2q}{2\pi r_F^2}\phi(\mathbf{q})e^{i|\mathbf{q}-\mathbf{x}|^2/2r_F^2} = -\int \frac{d^2q}{2\pi r_F^2}\phi(\mathbf{q}+\mathbf{x})e^{i|\mathbf{q}|^2/2r_F^2}$
  - where here $r_F = \sqrt{\lambda z/2\pi}$
- Consider a sinusoidal phase fluctuation: $\phi(\mathbf{q}) = \phi_0 \cos(\mathbf{k}\cdot\mathbf{q}) = \frac{1}{2}(\phi_0 e^{i\mathbf{k}\cdot\mathbf{q}} + \text{c.c.})$
- for which the observer plane ampliitude is
  - $\delta\psi(\mathbf{x}; z) = -\frac{1}{2}\phi_0 e^{i\mathbf{k}\cdot\mathbf{x}} \int \frac{d^2q}{2\pi r_F^2}\, e^{i\mathbf{k}\cdot\mathbf{q}}\, e^{i|\mathbf{q}|^2/2r_F^2} + \text{c.c.}$
- where the integral is the 2D Fourier transform of the Fresnel function
  - but as is easily shown (by completing the square), the Fourier transform of a normalised Gaussian $g(\mathbf{q}) = (2\pi\sigma^2)^{-1}e^{-|\mathbf{q}|^2/2\sigma^2}$ is another Gaussian:
    - $\tilde{g}(\mathbf{k}) = \int \frac{d^2q}{2\pi\sigma^2}\, e^{\mathbf{k}\cdot\mathbf{q}}\, e^{-|\mathbf{q}|^2/2\sigma^2} = e^{-\sigma^2|\mathbf{k}|^2/2}$
  - so, replacing $\sigma^2 \to ir_F^2 = i\lambda z/2\pi$, the desired integral is
    - $\int \frac{d^2q}{2\pi r_F^2}\, e^{i\mathbf{k}\cdot\mathbf{q}}\, e^{i|\mathbf{q}|^2/2r_F^2} = -ie^{ir_F^2|\mathbf{k}|^2/2} = -ie^{i\lambda z|\mathbf{k}|^2/4\pi}$
- so the perturbation to the complex amplitude is
  - $\delta\psi(\mathbf{x}; z) = ie^{i\lambda z|\mathbf{k}|^2/4\pi}\frac{1}{2}(\phi_0 e^{i\mathbf{k}\cdot\mathbf{x}} + \text{c.c.}) = i\phi(\mathbf{x})e^{i\lambda z|\mathbf{k}|^2/4\pi}$
  - or
  - $\boxed{\delta\psi(\mathbf{x}; z) = e^{i\lambda z|\mathbf{k}|^2/4\pi}\delta\psi(\mathbf{x}; z = 0)}$

- so the perturbation to the field amplitude, which starts off being purely imaginary, simply gets rotated on the Argand plane
  - so after a distance $z = 2\pi^2/\lambda|\mathbf{k}|^2$ the perturbation is $\delta\psi = -\phi_0 \cos(\mathbf{k}\cdot\mathbf{x})$ and is purely real
  - and, since the flux density is $F \propto |\psi|^2 = |1+\delta\psi|^2 \simeq 1+2\delta\psi$, we will see cosinusoidal flux density modulation on this plane $F(\mathbf{x}) \propto (1 - m\cos(\mathbf{k}\cdot\mathbf{x}))$ with so called *'modulation index'* $m = 2\phi_0$
  - and after a further distance $\Delta z = 2\pi^2/\lambda|\mathbf{k}|^2$ the perturbation is purely imaginary and has returned to being a pure phase error (but with opposite sign from the original)
  - aside: this phenomenon has important implications in *adaptive optics*
    * in an AO system the goal is to analyse some of the light coming into the detector to figure out the wave-front corrugations and then correct for this with a 'rubber' mirror
    * but this does not work if the phase errors imprinted by the atmosphere have 'rotated' into amplitude errors
      · which also cause image degradation
    * so what has to be done is to introduce some optics that creates, at some point in the beam, a 'conjugate' of the layer that was producing the seeing degradation (and thereby turn any amplitude errors back into correctable phase errors)
    * since the atmospheric seeing does not, in general, arise from a single layer, this leads to what are called 'multi-conjugate' AO systems
- the general initial perturbation to the wave amplitude $\delta\psi(\mathbf{x})$ can be expressed as a Fourier sum over waves with different spatial frequencies $\mathbf{k}$
  - for Kolmogorov turbulence the phase error $\phi(\mathbf{x})$ has power spectrum with $\langle\phi^2\rangle_k \sim k^2 P_\phi(k) \simeq (kr_{\rm diff})^{-5/3}$, which becomes large at small k
  - but the real part of the amplitude perturbation is equal to the phase error times $\sin\left(\lambda z k^2/4\pi\right) = \sin\left(k^2 r_{\rm F}^2/2\right)$
  - so the power spectrum of the flux density fluctuations is
    * $\langle(\delta F)^2\rangle_k \sim k^2 P_F(k) \simeq (kr_{\rm diff})^{-5/3}\sin^2(k^2 r_{\rm F}^2/2)$
  - it follows that $k^2 P_F(k) \propto k^{7/3}$ for $kr_{\rm F} \ll 1$ and $k^2 P_F(k) \propto k^{-5/3}$ for $kr_{\rm F} \gg 1$ and that the maximum modulation index will be at $k \sim 1/r_{\rm F}$ with $m \simeq (r_{\rm F}/r_{\rm diff})^{6/5}$ which, since we are in the weak scattering regime ($r_{\rm diff} \gg r_{\rm F}$) is small compared to unity (see figure 33 for a demonstration of this)
  - and the temporal power spectrum of the light curve of a source should show the same behaviour with $\omega = kv$ where v is the speed the observer is moving relative to the pattern of flux-density variations (which itself will, in general, be moving owing to the motion of the source and the scattering screen).
- another observable characteristic is that the modulation index will depend on wavelength
  - for a given 3D perturbation to the electron density n_e the perturbation to the refractive index δn scales as λ^2, as does the height δh of the wavefront corrugation,
  - so the phase error $\delta\phi = 2\pi\delta h/\lambda$ scales linearly with wavelength
  - so we should see a modulation index that has oscillations modulating an overall increasing trend of the index with λ

6.3 Scintillation in the strong scattering regime

- Increasing the strength of the refractive index fluctuations; the thickness of the scattering screen; the distance to the observer; and the observing wavelength all increase the scintillation strength.
- If $r_{\rm diff} \lesssim r_{\rm F}$ we are said to be in the *strong scattering regime*, and the character of the scintillation changes in a qualitative manner.

Table 1. *Examples of scattering media in astronomy, with typical values of* λ, D, r_F, *and* r_{diff}

medium	$\frac{\lambda}{\text{cm}}$	$\frac{D}{\text{cm}}$	$\frac{r_F}{\text{cm}}$	$\frac{r_{diff}}{\text{cm}}$	régime of scattering: weak	régime of scattering: strong
optical						
Earth's atmosphere	5×10^{-5}	10^6	3	10	mostly	near horizon
planetary atmospheres[a]	10^{-4}	10^{14}	4×10^4	10^2-10^6	early in occultation	deep in occultation
radio						
troposphere	20	10^5	6×10^2	$\sim10^5$	yes	no
ionosphere	3×10^2	3×10^7	4×10^4	$\sim10^5$	yes	sometimes
solar wind	10^2	10^{13}	10^7	$>10^7$	mostly	close to the Sun
interstellar medium	10^2	10^{21}	10^{11}	$\sim10^9$	no	yes

[a] Stars scintillate due to scattering in planetary atmospheres during occultations. The scattering is initially weak but becomes strong deep in the occultation (cf. Narayan & Hubbard 1988).

Figure 28: Some useful typical numbers for applications of scintillation in astronomy from Narayan (1992). Here λ is the observing wavelength; D the distance to the scattering screen; r_F is the Fresnel scale; and r_{diff} is the scale over which the phase fluctuations are of order 1 radian.

- one feature of the strong scattering regime is *'diffractive inter-stellar scintillation'* (DISS)
  - this can be understood as arising from interference between the radiation reaching the observer along different paths (*'multi-path propagation'*)
    - with the different paths have separation at the scattering screen, which called the *'refactive scale'*, $r_{ref} = r_F^2/r_{diff}$
  - this causes the flux density fluctuations on the observer plane – through which we are moving – to be 'speckly'
    - this gives flux density fluctuations that are always strong (for point sources at least) with modulation index unity.
    - and satisfy Rayleigh statistics (the flux density F has an exponential distribution $P(F)dF = \exp\big(-F/\overline{F}\big)dF/\overline{F}$)
  - the speckles on the observer plane have a length scale $\sim r_{diff}$ and therefore give rise to rapid fluctuations in flux-density
  - and they have a strong dependence on wavelength (different wavelengths having different speckle patterns)
  - multi-path propagation also results in observable *'pulse-broadening'* in pulsars (see figure 20)
- the other feature is *'refractive inter-stellar scintillation'* (RISS)
  - this is causes by the focusing of light by refractive index fluctuations on the scale r_{ref} and larger
  - this is largely understandable in terms of geometric optics
  - the fluctuations at scale r_{ref} – with flux density varying on the correspondingly long time-scale – are weaker than DISS
  - but these affect extended as well as point sources (though the effect is reduced if the source size is $\theta \gtrsim r_{ref}/D$)
- another feature that may possibly arise is the formation of 'caustics' by focussing
  - for this to happen would require a large 'inner scale' in order to suppress small scale corrugations in the wavefront
  - the formation of caustics from smooth wave-front deformations is illustrated in figure (32), which also discussed the transition between the wave- and geometric- or ray-optics descriptions
  - the scintillation caused by caustics – if indeed it happens – would be of large amplitude; appear at much longer time-scales than the speckle driven DISS; and would have a different dependence on observing wavelength

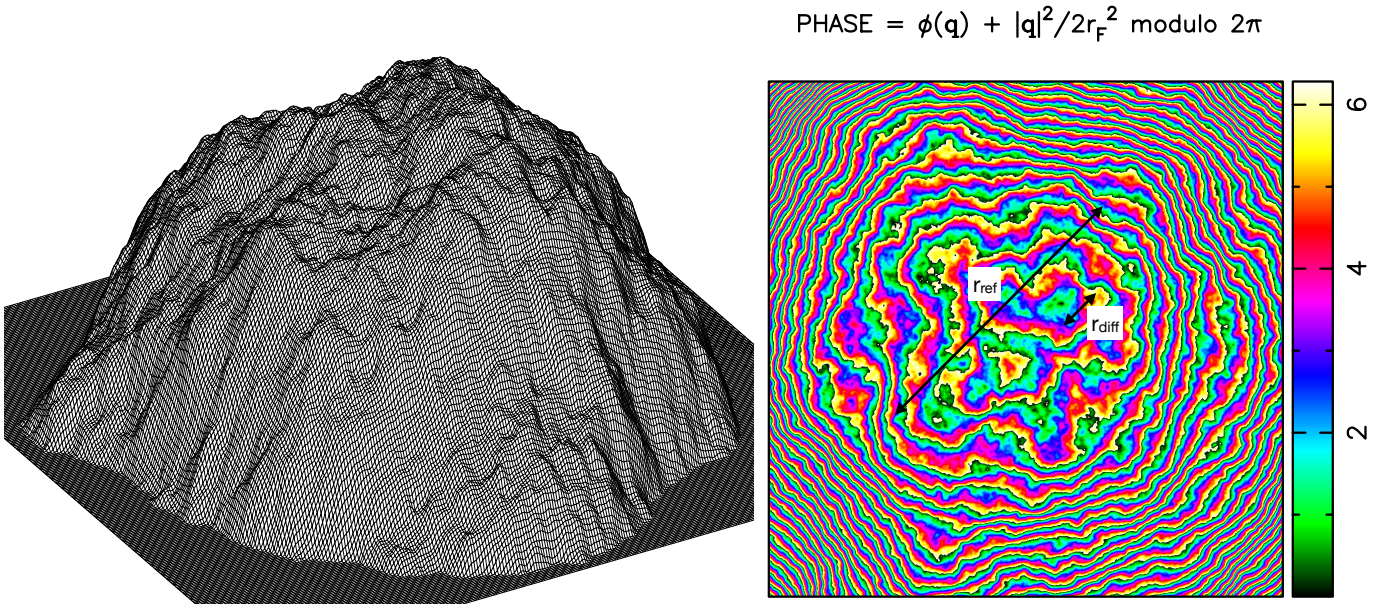


Figure 29: Illustration of DISS. The field amplitude ψ for some observer is the integral of $\exp\big(i(\phi(\mathbf{q}) + q^2/2r_F^2)\big)$, the complex exponential of the function plotted at left. Right shows the phase. On the edges of the 'mountain' the phase changes locally linearly and rapidly, so there is almost perfect cancellation. The main contribution to the ψ comes from regions of $\mathbf{q}$-space where the phase is close to stationary. In the strong scattering regime, the gradients of the Kolmogorov phase screen are sufficient to overwhelm the geometric phase-gradient out to a distance $r_{\rm ref} \sim r_{\rm F}^2/r_{\rm diff}$. The amplitude is the superposition of the contributions from patches which are individually of area $\sim r_{\rm diff}^2$, and within which there is little variation of phase. These are the multiple propagation paths and they lie within an area $\sim r_{\rm ref}^2$. These have different phases (different heights in the left plot) and, if one moves the observer, the phases change, which gives rise to the 'speckly' interference pattern on the observer plane (see figure 33).

6.4 Problems on turbulence, imaging and scintillation

6.4.1 Pulse dispersion

- assume below that one is observing radiation at frequencies well above the plasma frequency
- show that the arrival time delay of radiation that was emitted as a short transient pulse scales as wavelength squared and determine the constant of proportionality
- show that the amplitude of the waves – i.e. the envelope function – scales as $\omega^{3/2}$
  - hint: First, figure out what is the initial spectrum $P(k) = |\tilde{f}_{\mathbf{k}}|^2$ as a function of spatial frequency k.
  - then determine, for some time t, the range of distance Δx corresponding to some range of wavenumber Δk
  - Then use Parseval's theorem to relate $\int_x^{x+\Delta x} dx f^2(x)$ to the initial power in the corresponding range of spatial frequency.
- use the above to show that the amplitude falls with time delay as $(\Delta t)^{-3/4}$

6.4.2 Pulse broadening

- show that the pulse broadening from "multi-path propagation" should obey $\Delta t \propto \lambda^{22/5}$
- how does this theoretical prediction compare with the data shown in figure 20?
- using values the from Narayan's table reproduced above estimate the pulse-broadening (or 'scattering') time-scale. How does this compare with the data shown here (for the galactic centre magnetar SGR 1745-2900 whose period is about 3s)? How might this discrepancy be accounted for?

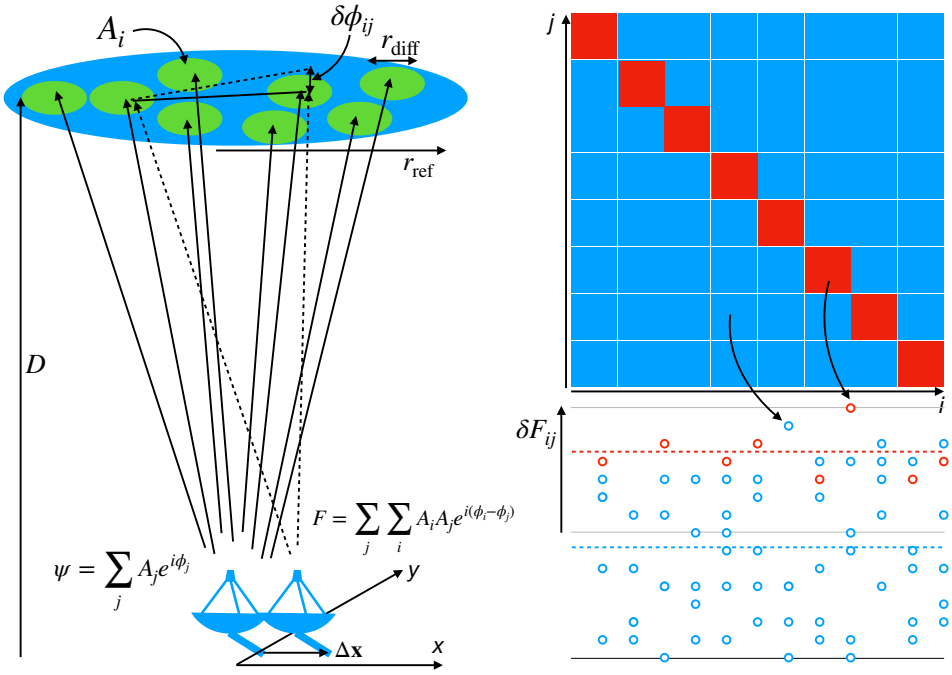


Figure 30: Simple model for why the speckle size in DISS is on the order of $r_{\rm diff}$. The field amplitude ψ is the sum of contributions from N patches of area A_i spread over the multi-path region of size $\sim r_{\rm ref}^2$, with phases ϕ_i having a perturbation from the refractive index fluctuations and also a geometric contribution depending on their location. The flux density $F = \psi\psi^\star$ is therefore a double sum with N diagonal elements (red) that have $\delta F_{ii} = A_i^2$ which is always positive and $N(N-1)$ off-diagonal elements (blue) that give contributions $\delta F_{ij} = A_i A_j e^{i(\phi_i - \phi_j)}$ that may be positive or negative (with $\delta F_{ij} + \delta F_{ji} = 2A_i A_j \cos(\phi_i - \phi_j)$). The 'root-N' RMS of the off-diagonal contributions is equal to the mean of the diagonal contributions but is equally likely to be positive or negative. If the observer moves a distance Δx this will change the phase differences of a pair by $\delta\phi_{ij} \sim (\Delta x/D) \times r_{\rm ref}/\lambda$. But $r_{\rm ref} \sim r_{\rm F}^2/r_{\rm diff}$, and $r_{\rm F}^2 \sim F\lambda$, so if $\Delta x \sim r_{\rm diff}$ the phase shift $\delta\phi_{ij}$ is roughly unity. Moving the observer by a distance $r_{\rm diff}$ is therefore sufficient to give a new realisation of the randomly fluctuating part of F; i.e. cause a significant change in F, hence the speckle size must be $\sim r_{\rm diff}$.

6.4.3 Timescale for scintillation caused by the ISM

- using values from Narayan's table reproduced above, and assuming relative motions of order $v \sim$ 100km/s, estimate the time-scale for strong scintillation of a typical pulsar
- how small would a source need to be to show strong scintillation in such circumstances?

6.4.4 2-point function of the EM field

This is a rather technical problem – but a good exercise nonetheless, particularly the latter parts which serve as a good preparation for the later cosmology section where Gaussian random fields play a pivotal role.

- Use the Fresnel-Kirchhoff equation (1) to show that the two point function of the complex field amplitude $\xi_\psi(\mathbf{d}) \equiv \langle \psi(\mathbf{x})\psi^\star(\mathbf{x}+\mathbf{d})\rangle_{\mathbf{x}}$ is independent of the observer-scattering screen distance D.
  - This may well seem surprising, given that the very character of the scintillation pattern depends critically on the distance (the flux density fluctuations being zero immediately down-stream of the screen, and the transition from weak- to strong-scattering being at a the distance where $r_{\rm diff}$, which depends only on the structure of the screen, and $r_{\rm F} \sim \sqrt{D\lambda}$, which increases with D, are the same.
  - the reason that $\xi_\psi(\mathbf{d})$ is inadequate to describe statistical properties of scintillation is that that would require e.g. the 2-point function of the flux-density F which, being the square of the field amplitude, involves the 4-point function of the phase.

- the invariance of $\xi_\psi(\mathbf{d})$ is important and useful as it is something that radio astronomers directly observe – it is the 'visibility' – and it relates directly to properties of point-spread function

- Show that if the phase error $\phi(\mathbf{r})$ is a statistically homogeneous Gaussian random field (as the central limit theorem encourages us to believe – see appendix) $\xi_\psi(r) = \exp(-(\xi_\phi(0) - \xi_\phi(r)))$, where $\xi_\phi(r) = \langle\phi(\mathbf{x})\phi(\mathbf{x}+\mathbf{r})\rangle_\mathbf{x}$.

- Hints:

  - For a Gaussian random field the probability distribution for the values of the field at two points $\mathbf{r}_1$ and $\mathbf{r}_2$ is $p(\phi_1,\phi_2)d\phi_1 d\phi_2 = (2\pi|M|)^{-1}d\phi_1 d\phi_2 \exp\Big(-\phi_i M_{ij}^{-1}\phi_j/2\Big)$ where $M_{ij} = \langle\phi_i\phi_j\rangle$ is the *'covariance matrix'* and M_{ij}^{-1} and $|M|$ are its inverse and determinant.
  - the above generalises to the n-point function $p(\phi_1,\phi_2,\ldots,\phi_n)$, but we will only need the 2-point function here
  - for a statistically homogeneous and isotropic random field the *'ensemble average'* $\langle\phi(\mathbf{r}_1)\phi(\mathbf{r}_2)\rangle$ i.e. the average over an ensemble of realisations (with $\mathbf{r}_1$ and $\mathbf{r}_2$ fixed) is the same as the spatial average $\langle\phi(\mathbf{x})\phi(\mathbf{x}+\mathbf{d})\rangle_\mathbf{x}$ with $\mathbf{d} = \mathbf{r}_1 - \mathbf{r}_2$

6.4.5 The PSF of images seen through a turbulent medium

- The 2-point function function $\xi_\psi(\mathbf{q})$ considered above plays a central role in imaging – it is also the quantity that is measured in aperture synthesis radio astronomy

- show that for a telescope with aperture diameter $D \gg r_{\rm diff}$ (or equivalently D much bigger than the Fried length) the point spread function $g(\mathbf{r})$ – time averaged for a long exposure – is the Fourier transform of $\xi_\psi(\mathbf{q})$

- if we observe a scene, for which the brightness $I(\hat{\mathbf{n}})$ can generally be written as a Fourier sum of plane-wave components, by what fraction are high-frequency components in the resulting image suppressed as compared to low-frequencies (that are well resolved)

- discuss why it is said that $\tilde{g}(\mathbf{k})$ defines the 'spatial bandwidth' of the imaging system

- using Kolmogorov scaling, show that, for a large telescope, the '2nd moment' of the PSF $\langle|\mathbf{r}|^2\rangle = \int d^2r r^2 g(r) / \int d^2r g(r)$ is ill defined (the integral in the numerator does not converge). Hint: establish and use the relationship between moments of $g(\mathbf{r})$ and derivatives of $\tilde{(}\mathbf{k} = 2\pi\mathbf{q}/L\lambda)$ at $\mathbf{k} = 0$.

- show that, in fact, the 'wings' of the PSF have a power law form: $g(r) \propto r^{-\gamma}$ and determine the index γ

- while the width $\langle|\mathbf{r}|^2\rangle$ as defined above is ill defined, show that an alternative measure of the width $r_{\rm eff} = (\nabla^2 g/g)^{-1/2}_{\mathbf{r}=0}$ is well defined

- discuss the effect of 'fine-scale mirror roughness' on the wings of the PSF

- discuss the relation between the various angular scales: diffraction, Fresnel, refraction, and telescope resolution and speckle size

- pulse-broadening is said to arise from "multi-path propagation". If one had a sufficiently high resolution telescope, would one resolve the separate images that this terminology suggests?

6.4.6 The structure function of scintillation speckles

- in the strong scattering regime one may assume that the real and imaginary parts of the field amplitude on the observer plane $\psi(\mathbf{x}) = a(\mathbf{x}) + ib(\mathbf{x})$ are independent Gaussian random fields with structure functions $\langle a(\mathbf{x})a(\mathbf{x}+\mathbf{d})\rangle_\mathbf{x} = \langle b(\mathbf{x})b(\mathbf{x}+\mathbf{d})\rangle_\mathbf{x} = \frac{1}{2}e^{-S_\phi(\mathbf{d})/2}$

- use this to show that the structure function of the flux density is

  - $S_F(\mathbf{d}) = \langle F(\mathbf{x})F(\mathbf{x}+\mathbf{d})\rangle_\mathbf{x} = 1 + \exp(-S_\phi(\mathbf{d}))$

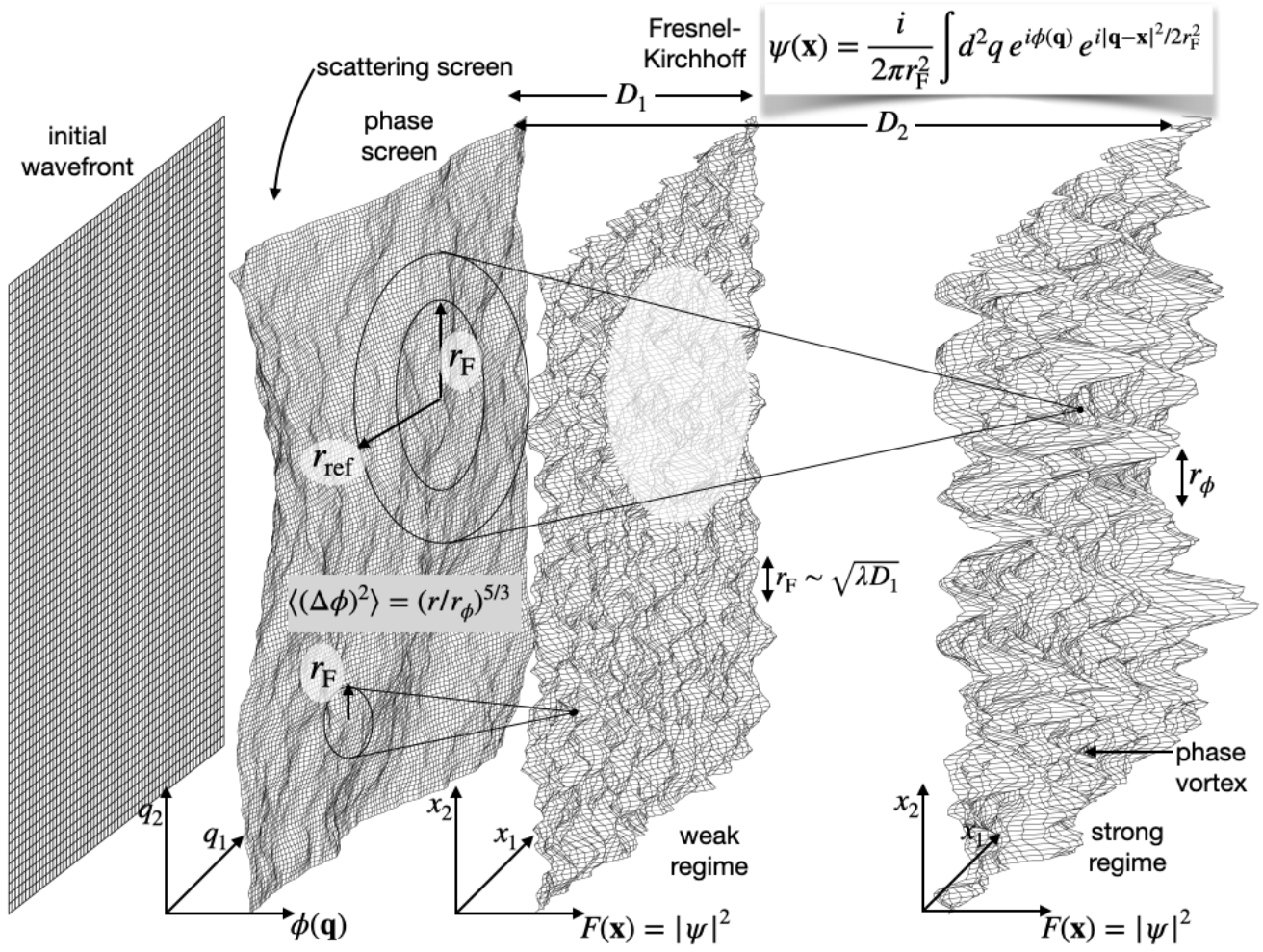


Figure 31: From left to right, the surfaces show an initially flat wave-front and then the corrugated wave-front – the 'phase-screen' that emerges after passing through some turbulent medium. For isotropic Kolmogorov turbulence this has structure function $S_\phi(d) = \langle(\phi_1 - \phi_2)^2\rangle = (d_{12}/r_\phi)^{5/3}$ were r_ϕ ($r_{\rm diff}$ in the text) is the separation for which the rms phase difference is 1-radian. The phase-screen here is a realisation of a Gaussian random field with $P_\phi(k) \propto k^{-11/3}$. The next surface shows the flux density $F(\mathbf{x})$ on the observed plane for a distance D such that the Fresnel scale $r_{\rm F} \sim \sqrt{\lambda D} \ll r_\phi$, which is the 'weak-scattering' regime. This is essentially a 'high-pass' filtered version of the phase-screen; the 'smoothing' with the Fresnel function $\exp(-|\mathbf{q}|^2/2r_{\rm F}^2)$ suppressing power at $\mathbf{k} \ll 1/r_{\rm F}$, so the dominant fluctuations appear at the Fresnel scale. Higher frequency modes undergo oscillation in the weak regime, 'rotating' from phase- to amplitude-errors and back again. The resulting power spectrum is $P_F(k) \propto k^{-11/3}\sin^2((kr_{\rm F})^2/2)$ which as zeros at frequencies such that the perturbation to wave amplitude is a pure phase shift. The rightmost surface shows the 'speckles' that form on the observer plane at a greater distance such that $r_{\rm F} \sim \sqrt{\lambda D} \gg r_\phi$, which is the 'strong-scattering' regime. In this regime, any point on the observer plane receives significant contributions to the wave amplitude ψ from a very large region on the scattering plane: $\Delta q \lesssim r_{\rm ref} \sim r_{\rm F}^2/r_\phi$. It follows – via the central limit theorem – that the real and imaginary parts of $\psi(\mathbf{x}) = a(\mathbf{x}) + ib(\mathbf{x})$ are statistically independent Gaussian random fields with structure functions $S_a(\mathbf{d}) = S_b(\mathbf{d}) = \frac{1}{2}S_\psi(\mathbf{d}) \equiv \langle\psi(\mathbf{x})\psi^\star(\mathbf{x}+\mathbf{d})\rangle_{\mathbf{x}} = \exp(-S_\phi(\mathbf{d})/2)$. From this one can show that the flux density $F(\mathbf{x}) = a^2(\mathbf{x}) + b^2(\mathbf{x})$ has an exponential distribution: $P(F)dF = \exp(-F)dF$ and has a structure function $S_F(\mathbf{d}) = 1 + \exp(-S_\phi(\mathbf{d}))$. The flux density vanishes at points – with observer plane density $n \sim 1/r_\phi^2$ – where both a and b vanish, and the phase θ of $\psi = |\psi|e^{i\theta}$ has 'screw-like' discontinuities (the phase wraps by 2π going around these points) known as 'phase-vortices'.

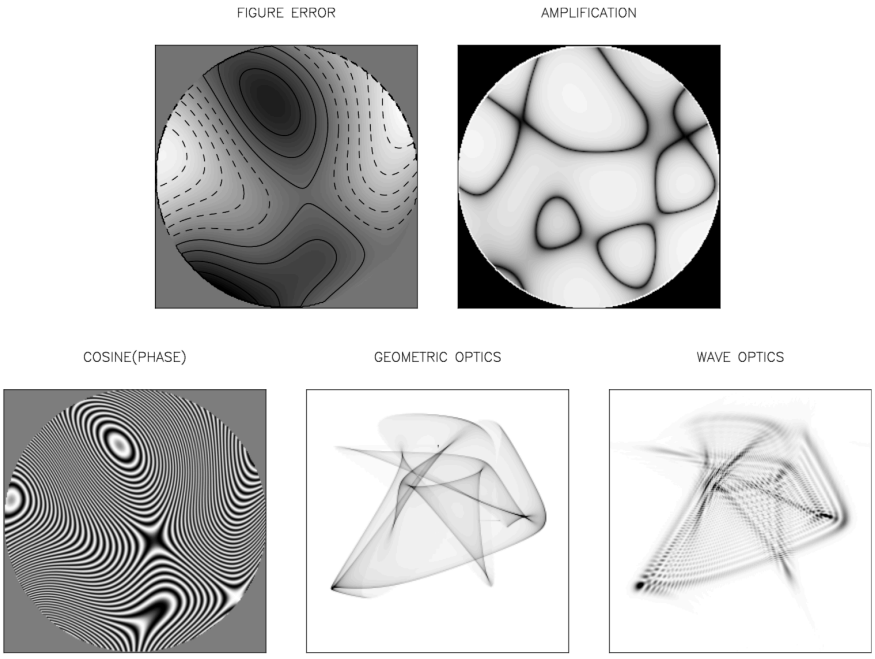


Figure 32: Transition between wave- and geometric- (or ray-) optics. This shows what happens in a telescope with some kind of aberration (the 'figure-error' – top left). This is the deformation of the wavefront, and dividing by $\lambda/2\pi$ gives the phase $\phi(\mathbf{q})$. The the Fresnel-Kirchhoff integral convolves the Fresnel function with $\exp(i\phi(\mathbf{q}))$, the real part of which is shown at the bottom left, and (squared) gives the flux density on the focal plane (bottom right). This shows the effect of multi-path interference. The bottom centre panel shows the geometric optics result which has caustics, the general form of which can be seen in the wave-optics image.

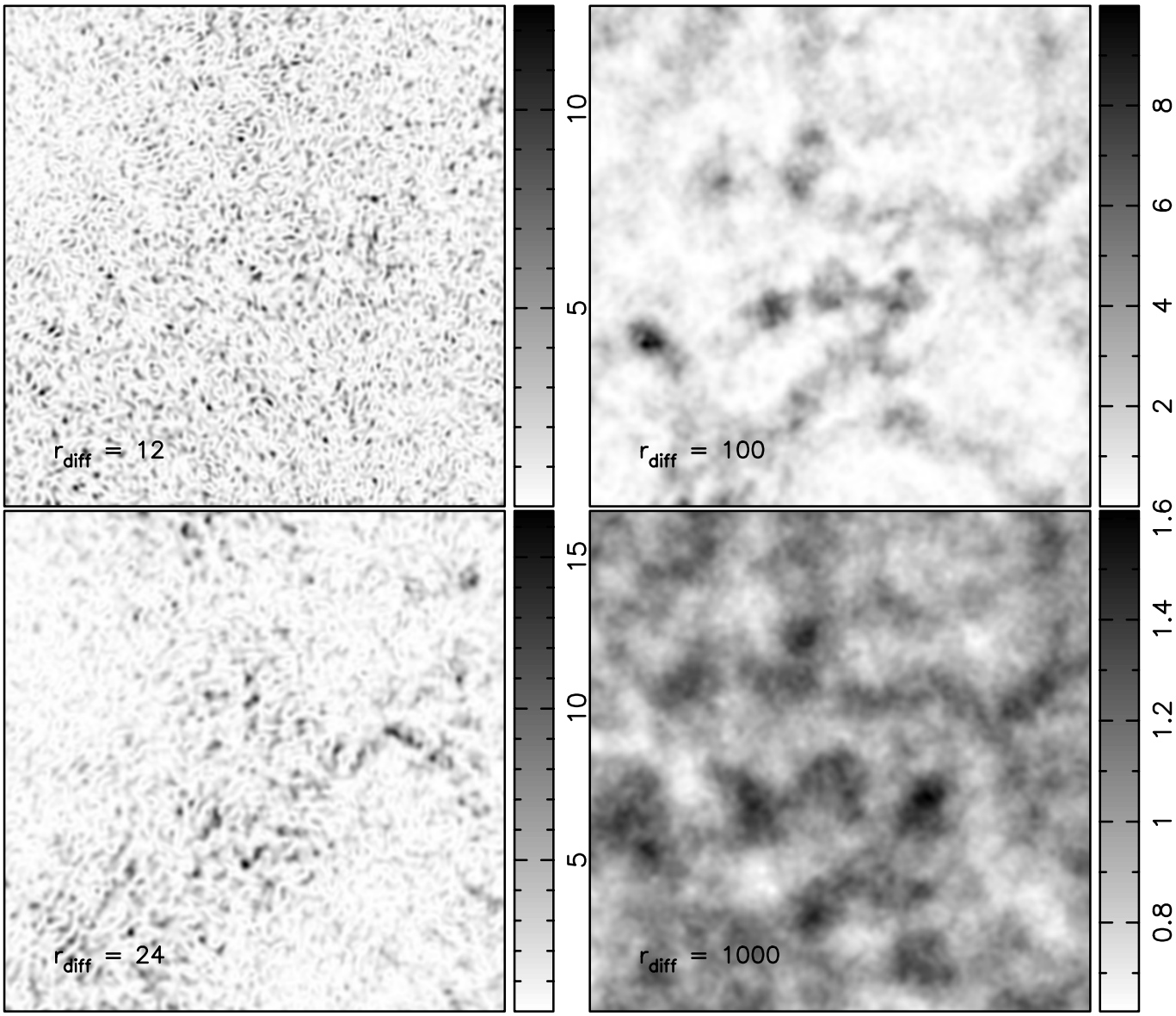


Figure 33: Observer-plane flux density maps generated by a Kolmogorov phase-screen for various values of the diffraction scale $r_{\rm diff}$. The Fresnel scale was fixed to be 100 pixels, and the size of the box size shown here is 20 times larger. These show how the character of the scintillation changes depending on the ratio $r_{\rm diff}/r_{\rm F}$. For $r_{\rm diff} \ll r_{\rm F}$ (upper left), we see speckles with size on the order of $r_{\rm diff}$; the modulation index is large; and are uniformly distributed; there is very little modulation on large scales. That's because the refractive scale $r_{\rm ref} \sim r_{\rm F}^2/r_{\rm diff}$ – being the size of regions on the scattering screen within which there is multi-path propagation – is very large for small $r_{\rm diff}$ and there is very little focussing on that scale (the focal length is much larger than the observer-scattering screen distance D). As we increase $r_{\rm diff}$ (lower left) the size of the speckles increases and $r_{\rm ref}$ decreases, so we start to see large scale flux density modulation caused by the focussing effect of the $r_{\rm ref}$-scale wavefront corrugations. At the other extreme (bottom right) where $r_{\rm diff} \gg r_{\rm F}$ and we are in the weak scattering regime the modulation is strongest on the scale $r_{\rm F}$, but is weak (note the scale bars at the right of the images).

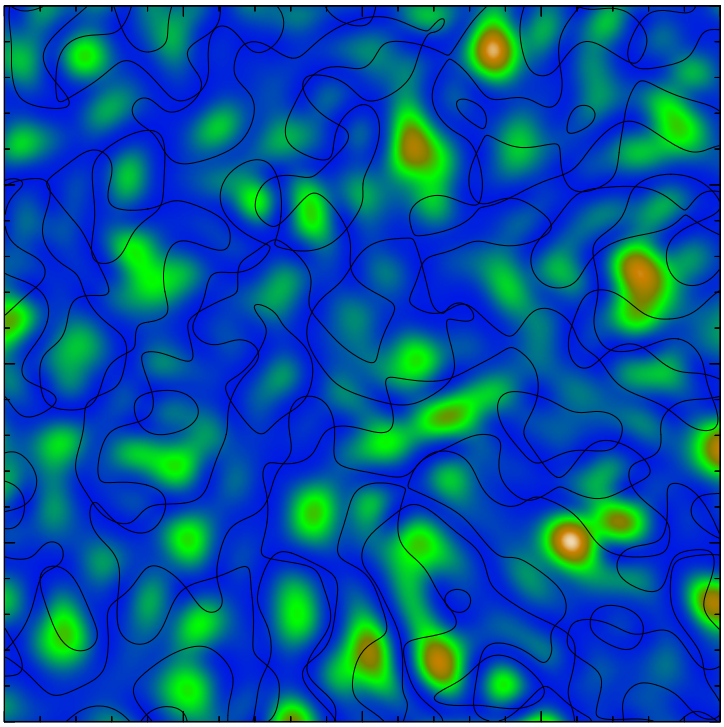


Figure 34: Blow up of part of the upper-left panel in figure (33) which depicts the strong scattering regime. The colour image is the 'speckly' flux-density $F(\mathbf{x}) = \psi\psi^\star$. Lines are where the real and imaginary parts of the field amplitude $\psi(\mathbf{x})$ vanish. Where they both vanish $F = \psi\psi^\star = 0$. At such dark points, the *phase* $\theta(\mathbf{x})$ of the field (defined by $\psi(\mathbf{x}) = |\psi|e^{i\theta}$) has a *screw-like discontinuity*; the phase wraps by 2π if one follows a path around a zero. So these are *'phase-vortices'*. In the strong scattering regime, the field $\psi(\mathbf{x})$ receives contributions from many patches (of size $\sim r_{\rm diff}$; the phase correlation length) and the central limit theorem tells us that the field becomes a (complex) Gaussian random field. It is fully determined by the auto-correlation function $\xi_\psi(\mathbf{d}) \equiv \langle\psi(\mathbf{x})\psi^\star(\mathbf{x}+\mathbf{d})\rangle_{\mathbf{x}} = e^{-S_\phi(d)/2}$, where $S_\phi(d)$ is the phase structure functon or, equivalently, by the power spectrum $P_\psi(\mathbf{k}) = |\tilde{\psi}_{\mathbf{k}}\tilde{\psi}^\star_{\mathbf{k}}|$. This explains the universal character of strong scintillation.

A Statistics of random fields

A.1 Discrete, continuous and semi-continuous Fourier transforms

A.1.1 The discrete Fourier transform in 1D

- given an array F_Q where $Q = \{0, 1, 2 \ldots, N-1\}$ is an integer index we define the discrete transform
  - $$\boxed{\tilde{F}_K = \sum_{Q=0}^{N-1} F_Q e^{2\pi i K Q/N}}$$
- for $K = \{0, 1, 2 \ldots, N-1\}$
- from which we find
  - $\sum_K \tilde{F}_K e^{-i2\pi K Q/N} = \sum_{Q'} F_{Q'} \sum_K e^{-2\pi i K(Q-Q')/N} = \sum_{Q'} F_{Q'} N \delta_{QQ'} = N F_Q$
- where the fact that e.g. $\mathrm{Re}(\sum_K e^{-2\pi i K(Q-Q')/N}) = \sum_K \cos(-2\pi i K \Delta Q/N)$ vanishes unless $\Delta Q = Q - Q' = 0$, and similarly for the imaginary part, is most obvious – by symmetry – when N is a power of 2, but is true in general
- so the inverse transform is
  - $$\boxed{F_Q = N^{-1} \sum_{K=0}^{N-1} \tilde{F}_K e^{-2\pi i K Q/N}}$$

A.1.2 The Nyquist frequency and symmetry of the discrete transform

- $\tilde{F}_{K=0}$ is the amplitude of the 'DC mode'; $\tilde{F}_{K=1}$ is that of the 'fundamental frequency', where 1 wave fits in the length of the array, and the others are modes of increasing spatial frequency
- but only up to $K = N/2$, which is known as the *'Nyquist frequency'* K_N
  - for which the phase factors $e^{2\pi i K_\mathrm{N} Q/N} = e^{i\pi Q} = \{1, -1, 1, -1, \ldots\}$ are real and simply oscillate at the maximum allowed frequency
- higher values of K have phase factors that oscillate at lower frequencies
  - this phenonenon being known as 'aliasing'
- going down to $K = N - 1$ which oscillates at the fundamental frequency
- the phase-factor $e^{2\pi i K Q/N}$ is symmetric about K_N: $e^{2\pi i (K_\mathrm{N}+M)Q/N} = (e^{2\pi i (K_\mathrm{N}-M)Q/N})^\star$
- from which it follows that, if F_Q is real, $\tilde{F}_{K_\mathrm{N}+M} = \tilde{F}^\star_{K_\mathrm{N}-M}$
  - so $\tilde{F}_K$, despite having N complex components is fully determined by N real numbers
- it is best to mentally 'wrap' the array $\tilde{F}_K$, and think of K having negative as well as positive frequencies that run from $K = -(K_\mathrm{N} - 1)$ to $K = +(K_\mathrm{N} - 1)$ with zero-frequency in the centre
  - and with the Nyquist frequency component $\tilde{F}_{K_\mathrm{N}}$ at one or other end
  - as it is of neither positive or negative frequency

A.1.3 The discrete Fourier transform in n-dimensions

- this generalises to n-dimensions (e.g. for a $N \times N$ grid in 2-dimensions)
  - $\boxed{\tilde{F}_{\mathbf{K}} = \sum_{\mathbf{Q}} F_{\mathbf{Q}} e^{2\pi i \mathbf{K}\cdot\mathbf{Q}/N}} \quad \Leftrightarrow \quad \boxed{F_{\mathbf{Q}} = N^{-n} \sum_{\mathbf{K}} \tilde{F}_{\mathbf{K}} e^{-2\pi i \mathbf{K}\cdot\mathbf{Q}/N}}$
- where one should think of $\mathbf{K}$ values as living on a cubical lattice with DC mode $\mathbf{K} = \mathbf{0}$ (more or less) in the centre
- the asymmetry arising because the Nyquist frequency planes have to be put at either positive or negative values

A.1.4 The continuous Fourier transform

- to make the transition to a continuous field, we let
  - $\mathbf{q} \equiv \epsilon \mathbf{Q}$
  - where ϵ is the physical 'pixel size', and
  - $\mathbf{k} \equiv 2\pi \mathbf{K}/N\epsilon = 2\pi \mathbf{K}/L$
  - where $L = N\epsilon$ is the size of the box
  - so $2\pi \mathbf{Q}\cdot\mathbf{K}/N = \mathbf{q}\cdot\mathbf{k}$
  - and we define
  - $f(\mathbf{q}) \equiv F_{\mathbf{Q}}$,
  - and
  - $\tilde{f}(\mathbf{k}) = \epsilon^n \tilde{F}_{\mathbf{K}}$
- in terms of which the forward transform is
  - $\tilde{f}(\mathbf{k}) = \epsilon^n \tilde{F}_{\mathbf{K}} = \sum_{\mathbf{Q}} \epsilon^n F_{\mathbf{Q}} e^{2\pi i \mathbf{K}\cdot\mathbf{Q}/N} \quad \Rightarrow \quad \int d^n q f(\mathbf{q}) e^{i\mathbf{k}\cdot\mathbf{q}}$
  - since the volume of a 'voxel' is ϵ^n so we can replace $\sum_{\mathbf{Q}} \epsilon^n \ldots \Rightarrow \int d^n q \ldots$
- and the inverse transform is
  - $f(\mathbf{q}) = F_{\mathbf{Q}} = \frac{1}{(2\pi)^n} \sum_{\mathbf{K}} \left(\frac{2\pi}{\epsilon N}\right)^n \epsilon^n F_{\mathbf{K}} e^{-i\mathbf{k}\cdot\mathbf{q}} \quad \Rightarrow \quad \int \frac{d^n k}{(2\pi)^n} f(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{q}}$
  - since the spacing in $\mathbf{k}$-space is $\Delta k = 2\pi/L = 2\pi/(\epsilon N)$ and so $\sum_{\mathbf{K}} (\Delta k)^n \ldots \Rightarrow \int d^n k \ldots$
- so, to summarise, for a continuous field $f(\mathbf{q})$ we can define the forward and inverse transforms as
  - $\boxed{f(\mathbf{q}) = \int \frac{d^n k}{(2\pi)^n} \tilde{f}(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{q}}} \quad \Leftrightarrow \quad \boxed{\tilde{f}(\mathbf{k}) = \int d^n q f(\mathbf{q}) e^{i\mathbf{k}\cdot\mathbf{q}}}$
- where we note that $f(\mathbf{q})$ and $\tilde{f}(\mathbf{k})$ have different units
  - if $f(\mathbf{q})$ is dimensionless then $\tilde{f}$ has units of (length)^n

A.2 The Dirac δ-function

- in the discrete case we used $\sum_K e^{-2\pi i K(Q-Q')/N} = N\delta_{QQ'}$
  - whose generalisation to n-dimensions is
  - $\boxed{\delta_{\mathbf{QQ'}} = N^{-n} \sum_{\mathbf{K}} e^{-2\pi i \mathbf{K}\cdot(\mathbf{Q}-\mathbf{Q'})/N}}$
  - i.e. $\delta_{\mathbf{QQ'}}$ is the inverse transform of $\tilde{F}_{\mathbf{K}} = 1$ (evaluated at $\mathbf{Q} - \mathbf{Q'}$)
- in the continuous case we define

- $\boxed{\delta(\mathbf{q}-\mathbf{q}') \equiv \epsilon^{-n}\delta_{\mathbf{Q}\mathbf{Q}'}}$

- which has the property that
  - $\boxed{\int d^n q' f(\mathbf{q}')\delta(\mathbf{q}-\mathbf{q}') = f(\mathbf{q})}$
  - since
  - $\int d^n q' f(\mathbf{q}')\delta(\mathbf{q}-\mathbf{q}') \Rightarrow (\sum_{\mathbf{Q}'}\epsilon^n)F_{\mathbf{Q}'}(\epsilon^{-n}\delta_{\mathbf{Q}\mathbf{Q}'}) = F_{\mathbf{Q}} \Rightarrow f(\mathbf{q})$

- and is the inverse transform of $\tilde{f}(\mathbf{k}) = 1$, since (using $\Delta k = 2\pi/\epsilon N$)
  - $\epsilon^{-n}\delta_{\mathbf{Q}\mathbf{Q}'} = \frac{1}{(\epsilon N)^n}\sum_{\mathbf{K}} e^{-2\pi i\mathbf{K}\cdot(\mathbf{Q}-\mathbf{Q}')/N} = (2\pi)^{-n}\sum_{\mathbf{K}}(\Delta k)^n e^{-\mathbf{k}\cdot(\mathbf{q}-\mathbf{q}')} \Rightarrow \int \frac{d^n k}{(2\pi)^n} e^{-i\mathbf{k}\cdot(\mathbf{q}-\mathbf{q}')}$

- so if whenever we see something like $\int \frac{d^n k}{(2\pi)^n} e^{-i\mathbf{k}\cdot(\ldots)}$ we can replace it by $\delta(\ldots)$.

- there are other representations of the Dirac δ-function
  - one is as a limit of a normalised Gaussian $\delta(\mathbf{q}) = \lim_{\sigma\to 0}(2\pi\sigma^2)^{-n/2}\exp\bigl(-|\mathbf{q}|^2/2\sigma^2\bigr)$
  - which has transform $\tilde{\delta}(\mathbf{k}) = \exp\bigl(-\sigma^2|\mathbf{k}|^2/2\bigr)$
  - which is approximately unity for $|\mathbf{k}| \ll 1/\sigma$

A.2.1 The semi-continuous Fourier transform

- all of the above is conventional
  - though some people like to 'symmetrise' the transforms by using $d^n q/(2\pi)^{n/2}$ and $d^n k/(2\pi)^{n/2}$
  - which is bad form

- for most fields we are concerned with it is very reasonable to think about the field as a continuous function $f(\mathbf{q})$

- and, if we were dealing with a field $f(\mathbf{q})$ with 'compact support',
  - i.e. non-zero only within some volume of linear size l say
  - assumed to be much smaller than the size of the fictitious periodicity volume L

- then $\tilde{f}(\mathbf{k})$ is also a continuous field
  - since the compactness of the support introduces a *coherence* in $\tilde{f}(\mathbf{k})$
  - and it becomes effectively continuous if sampled on a grid with $\Delta k \ll 1/l$

- but for the statistically homogeneous random fields – with infinite support — that we are interested in here (see examples in figure 35) that is not the case
  - for these the transform $\tilde{f}(\mathbf{k})$ is always 'grainy' at a fine scale
  - that scale being set by the fictitious volume size L

- this motivates an alternative – though equivalent – hybrid 'semi-continuous' formalism which proves to be useful, and in which we define $\tilde{f}_{\mathbf{k}} \equiv L^{-n}\tilde{f}(\mathbf{k})$, defined at discrete values of $\mathbf{k}$, in terms of which the transform and inverse transform pair is
  - $\boxed{f(\mathbf{q}) = \sum_{\mathbf{k}} \tilde{f}_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{q}}} \Leftrightarrow \boxed{\tilde{f}_{\mathbf{k}} = L^{-n}\int d^n q f(\mathbf{q}) e^{i\mathbf{k}\cdot\mathbf{q}}}$

- and we note that, in this convention, $\tilde{f}_{\mathbf{k}}$ and $f(\mathbf{q})$ have the same units

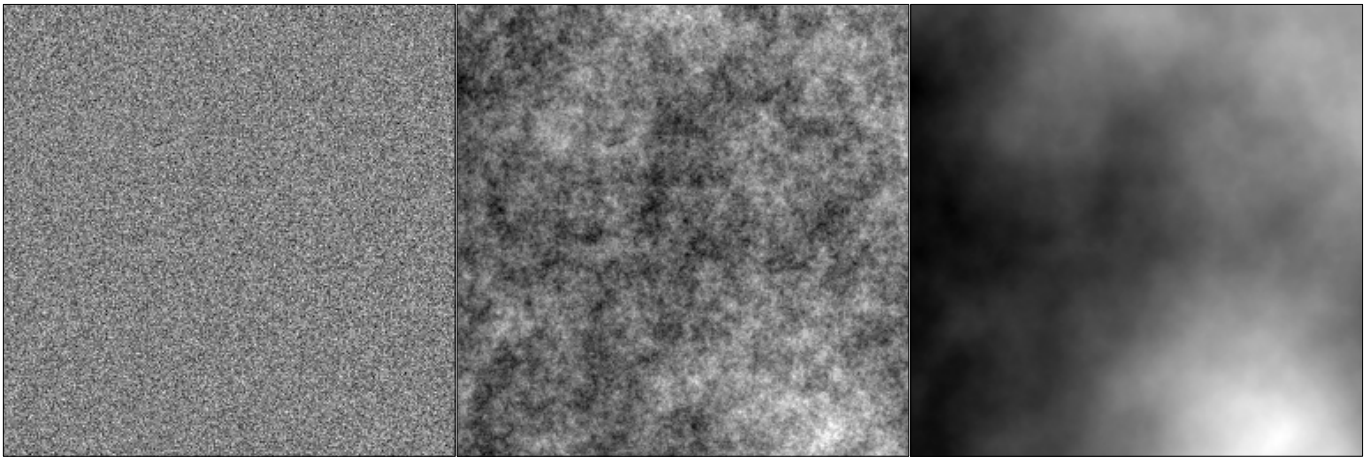


Figure 35: Some realisations of Gaussian random fields. They were generated by a creating 2D grid of complex $\tilde{f}_{\mathbf{k}}$ values with real and imaginary parts drawn from 1D Gaussian distributions (though the result is insensitive to the actual distribution used) with width given by $\sqrt{P(k)}$ and then performing the inverse transform to generate the images. In each case $P(k)$ was a power law $P(k) \propto k^n$. The left panel is for a flat, or 'white-noise' spectrum: $n = 0$. The autocorrelation function for this vanishes except at zero lag. If one were to smooth this by taking a moving average with a window of size L containing N pixels, the RMS fluctuation Δ_L would scale as $1/\sqrt{N}$, or equivalently as $\Delta \propto \sqrt{(k^2 P(k))_{k\sim 1/L}} \propto 1/L$. The right panel is for a very 'red' spectrum with $n = -4$. Here $\Delta_L \propto \sqrt{(k^2 P(k))_{k\sim 1/L}} \propto L$. The centre panel is for $n = -2$. Here $\Delta_L \propto \sqrt{(k^2 P(k))_{k\sim 1/L}}$ is independent of smoothing scale; so there is roughly the same variance on all scales. The structure function $S(d)$ in this case increases logarithmically with separation d. For the right hand image, $S(d) \propto d^2$. This is quite similar to the phase fluctuations imposed in wavefronts of light that has propagated through a turbulent atmosphere or the ISM, for which $S(d) \propto d^{5/3}$ corresponding to spectral index $n = -11/3$.

A.3 Autocorrelation function, power spectrum and the Wiener Khinchin theorem

- we'll work in 2-dimensions (and the field will be $f(\mathbf{q})$, and you might think of the phase error induced by a scattering screen as a concrete example), but the results are readily generalisable
- we'll imagine the field is periodic in a large square area $A = L^2$
- so we can write $f(\mathbf{q})$ as a Fourier series
  - $f(\mathbf{q}) = \sum_{\mathbf{k}} \tilde{f}_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{q}}$
  - where the $\mathbf{k}$-vectors lie on a cubical grid of spacing $2\pi/L$ and use
  - $\tilde{f}_{\mathbf{k}} = A^{-1} \int d^2q f(\mathbf{q}) e^{-i\mathbf{k}\cdot\mathbf{q}}$
  - from which we can see that $\tilde{f}_{-\mathbf{k}} = \tilde{f}^{\star}_{\mathbf{k}}$, so $\tilde{\phi}_{\mathbf{k}}$ has mirror symmetry
    - thus, we only specify the complex coefficients in half of the $\mathbf{k}$-plane
    - the other half following by symmetry
    - and it follows that if the field is 'pixelised' on a grid the number of real numbers needed to determine $f_{\mathbf{q}}$ are the same as those needed to specify $\tilde{f}_{\mathbf{k}}$
- we define the *autocorrelation function* at 'lag' $\mathbf{d}$ to be
  - $$\boxed{\xi_f(\mathbf{d}) \equiv \langle f(\mathbf{q}) f(\mathbf{q}+\mathbf{d}) \rangle_{\mathbf{q}} = A^{-1} \int d^2q f(\mathbf{q}) f(\mathbf{q}+\mathbf{d})}$$
  - or, writing $f(\mathbf{q}) = \sum_{\mathbf{k}} \tilde{f}_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{q}}$ and $f(\mathbf{q}+\mathbf{d}) = \sum_{\mathbf{k}} \tilde{f}^{\star}_{\mathbf{k}} e^{-i\mathbf{k}\cdot(\mathbf{q}+\mathbf{d})}$
  - $\xi_f(\mathbf{d}) = \sum_{\mathbf{k}} \sum_{\mathbf{k}'} \tilde{f}_{\mathbf{k}} \tilde{f}^{\star}_{\mathbf{k}'} A^{-1} \int d^2q e^{i(\mathbf{k}\cdot\mathbf{q} - \mathbf{k}'\cdot(\mathbf{q}+\mathbf{d}))} = \sum_{\mathbf{k}} \sum_{\mathbf{k}'} \tilde{f}_{\mathbf{k}} \tilde{f}^{\star}_{\mathbf{k}'} e^{-i\mathbf{k}'\cdot\mathbf{d}} \delta_{\mathbf{k}\mathbf{k}'}$
  - so

- $\boxed{\xi_f(\mathbf{d}) = \sum_{\mathbf{k}} |\tilde{f}_{\mathbf{k}}|^2 e^{-i\mathbf{k}\cdot\mathbf{d}}}$
- which is the *Wiener-Khinchin theorem*, which states that $\xi_f(\mathbf{d})$ is the inverse Fourier transform of the *power spectrum*
- $\boxed{P_f(\mathbf{k}) \equiv |\tilde{f}_{\mathbf{k}}|^2}$

- note that we are defining $\xi_f(\mathbf{d})$ as a *spatial average* (integrating over all pairs of points with separation $\mathbf{d}$ and dividing by the area A)

- an alternative would be to define $\xi_f(\mathbf{d})$ as an *ensemble average*
  - where we might set $\mathbf{q} = \mathbf{0}$ and average, over an ensemble of *realisations*, $f(\mathbf{0})f(\mathbf{d})$
  - in which case the W-K theorem would involve the ensemble average $\langle |\tilde{f}_{\mathbf{k}}|^2 \rangle$
  - which looks rather different: it might be some smooth function of $\mathbf{k}$, whereas $|\tilde{f}_{\mathbf{k}}|^2$ is, like $\tilde{f}_{\mathbf{k}}$ 'grainy' at a fine scale
  - and one might worry that one then needs to define some kind of statistical model for how the Fourier amplitudes $\tilde{f}_{\mathbf{k}}$ are generates
  - but the distinction is not really important, at least if we are dealing with fields (or 'processes' in the jargon) for which things like $\xi_f(\mathbf{d})$ tend to some sensible limit at we let the averaging volume $A \to \infty$
  - and this is indeed what we expect since, despite the fact that $|\tilde{f}_{\mathbf{k}}|^2$ is grainy, when we use the equation above to evaluate $\xi_f(\mathbf{d})$ we are summing over $|\tilde{f}_{\mathbf{k}}|^2$ times a smoothly varying function (of $\mathbf{k}$) $e^{i\mathbf{k}\cdot\mathbf{d}}$, so all that influences the outcome is the local average of $|\tilde{f}_{\mathbf{k}}|^2$

- the real distinction between the two types of power spectra arises as follows:
  - one might use a model for the ensemble average power spectrum $P_f(\mathbf{k})$
    - some smooth function of $\mathbf{k}$
    - for example here we have been using Kolmogorov's prediction that $P_\phi(\mathbf{k})$ is a power law with index $-11/3$
  - to generate *realisations* of the field $f(\mathbf{q})$ (as done to make figure 35)
    - a common assumption being that the real and imaginary parts of $\tilde{f}_{\mathbf{k}}$ are drawn from independent zero-mean Gaussian distributions with variance $P_\phi(\mathbf{k})/2$
    - but, as we have just discussed, this does not influence quantities like $\xi_f(\mathbf{d})$, nor does it affect higher order correlation functions
  - then, given a realisation, or given some data, one can measure the actual power spectrum $|\tilde{f}_{\mathbf{k}}|^2$
    - this will have some overall smooth variation – just like the ensemble – but with fluctuating graininess modulating this
    - for example, for the simulated data generated as above, $|\tilde{f}_{\mathbf{k}}|^2$ would have an exponential distribution (χ^2 with 2 'degrees of freedom')
  - and one might then do some sort of local smoothing to estimate the ensemble average $P_f(\mathbf{k})$
  - or one might generate, from $|\tilde{f}_{\mathbf{k}}|^2$, the 2-point function $\xi_f(\mathbf{d})$
  - and if one were doing this on a computer-generated realisation, the information in the graininess of $|\tilde{f}_{\mathbf{k}}|^2$ would not necessarily be lost.
    - This graininess is high frequency structure in $\mathbf{k}$-space, that would be encoded in what is effectively noise in $\xi_f(\mathbf{d})$ at large scales, where the ensemble average of ξ_f is essentially zero.
    - Aside from problems with round-off error, one could recover the full grainy $|\tilde{f}_{\mathbf{k}}|^2$.
    - One would not, however, be able to recover from this the individual components $\mathrm{Re}(\tilde{f}_{\mathbf{k}})$ and $\mathrm{Im}(\tilde{f}_{\mathbf{k}})$. We say that the *'phase-information* has been lost in extracting only $|\tilde{f}_{\mathbf{k}}|^2$

A.4 Gaussian random fields

A.4.1 The central limit theorem

- the *central limit theorem* says that if we have a set of N random variates x_i drawn independently from some probability distribution $p_x(x)dx$ (which we will assume has zero mean) and we make the sum of these: $X = \sum_i x_i$, then X will, for large N, tend to have a Gaussian distribution:
  - $P_X(X)dX = dX(2\pi\sigma_X^2)^{-1/2}\exp\bigl(-X^2/2\sigma_X^2\bigr)$
  - where the variance σ_X^2 (which is equal to $\int dX X^2 P_X(X)$) is given by $\sigma_X^2 = N\sigma_x^2$ with $\sigma_x^2 = \int dx x^2 p_x(x)$, *largely independent of the detailed form of* $p_x(x)$
- the proof of this is straightforward:
  - The distribution $P_X(X)$ is the multiple convolution of N factors $p_x(x)$
    - for example, for $N = 2$, $P(X) = P_2(X) = \int dy p_x(y) p_x(X - y) = p_x \otimes p_x$
    - and if we add a third variate, we convolve P_2 with another p_x and so on
  - the Fourier transform $\tilde{P}_X(\omega)$ of $P_X(X)$ – called the 'generating function' – is therefore, by virtue of the convolution theorem, the product of N-factors $\tilde{p}_x(\omega)$: $\tilde{P}_X(\omega) = \tilde{p}_x(\omega)^N$
  - if N is very large then P_X will be very broad compared to p_x
  - so $\tilde{P}_X(\omega)$ will be very narrow compared to $\tilde{p}_x(\omega)$
  - so, for ω-values of interest, we can expand the exponential in $\tilde{p}_x(\omega)$:
    - $\tilde{p}_x(\omega) = \int dx p_x(x) e^{i\omega x} \simeq \int dx p_x(x)(1 + i\omega x - \omega^2 x^2/2 \ldots) = 1 + i\omega\langle x\rangle - \omega^2\langle x^2\rangle/2 + \ldots$
    - but, by assumption, $\langle x\rangle = 0$ so
    - $\tilde{p}_x(\omega) \simeq 1 - \omega^2\sigma_x^2/2$
    - and taking the Nth power of this and using $\exp(y) = \lim_{N\to\infty}(1 + y/N)^N$ gives
    - $\tilde{P}_X(\omega) \simeq (1 - \omega^2\sigma_x^2/2)^N = \exp\bigl(-N\sigma_x^2\omega^2/2\bigr)$
    - whose inverse transform is another Gaussian
    - $P_X(X) \propto \exp\bigl(-X^2/2N\sigma_x^2\bigr)$

A.4.2 Gaussian n-point functions

- similar reasoning may be applicable to random fields. crudely speaking:
  - since $f(\mathbf{q}) = \sum_{\mathbf{k}} \tilde{f}_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{q}}$ is the sum of a very large number of random variates
  - then its value at a point should have a Gaussian probability distribution
  - and the joint distribution of the values of the field at a set of points should have a multi-variate Gaussian distribition
- there may be good physical reasons for this (as, perhaps, in cosmology, where it may stem from initial conditions)
- but in atmospheric or ISM physics this is perhaps more questionable
  - the phase-errors behind a thin screen of turbulent ISM or atmosphere is unlikely to be accurately Gaussian – even though it might well be reasonably statistically homogeneous in nature
  - but a *thick* slab in projection probably is – because we are adding the effect of many eddies
  - and it is more likely that the predictions of 'fully developed turbulence' for the structure function or power spectrum would apply
  - but in the strong scattering regime, even for a thin screen, the field ψ gets contributions from a large area of the screen, so the central limit theorem should apply
  - but it may be that the structure function of the phase fluctuations is different from the Kolmogorov predictions – as the 'outer scale' might be comparable to the scales one is sensitive to

- there are several types of interesting questions one can address with the Gaussian model (with ISM related applications in parentheses):
  - what is the joint distribution of the values of a field at a number of different places (e.g. the phase ϕ or the EM field amplitude ψ)?
  - what are the statistical properties of non-linear functions of a Gaussian random field (for example, $e^{i\phi(\mathbf{q})}$ appearing in the Fresnel-Kirchhoff integral for ψ, or flux density $F = |\psi|^2$)
  - for instance, in the ISM it would be nice to be able to predict the properties of light curves of scintillating sources in order to address questions like:
    * what is the probability distribution for the flux-density?
    * what is the structure function for the flux density?
    * how frequent are peaks in light curves? or other ways to characterise the size of speckles?
  - with the goal being to relate these observable quantities to the structure functions of the ISM
- The multi-variate formalism proceeds as follows:
  - consider some 'vector' of m values of the field $f(\mathbf{q})$ at a set of points:
    * $\mathbf{f} \equiv \{f_1, f_2, \ldots\} = \{f(\mathbf{q}_1), f(\mathbf{q}_2), \ldots\}$
  - what is the probability distribution $P(\mathbf{f})d^m f$ for the vector $\mathbf{f}$?
    * where we can think of this as the distribution over an ensemble of realisations
    * or as a spatial average where we sample all possible values for $\mathbf{q}_1$ with the other $\mathbf{q}_m - \mathbf{q}_1$ held fixed
  - each of the f_m is a sum: $f_m = \sum_{\mathbf{k}} f_{\mathbf{k}m} = \sum_{\mathbf{k}} \tilde{f}_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{q}_m}$ so the central limit says they will have a Gaussian probability distribution
    * $\boxed{P(\mathbf{f})d^m f = d^m f (2\pi|M|)^{-1/2} \exp\Big(-f_i M^{-1}_{ij} f_j/2\Big)}$
    * where the covariance matrix $M_{ij} = \langle f_i f_j \rangle = \sum_{\mathbf{k}} m_{\mathbf{k}ij}$
    * with $m_{\mathbf{k}ij} = \langle f_{\mathbf{k}i} f_{\mathbf{k}j} \rangle$ being the covariance matrix for the individual $\mathbf{f}_{\mathbf{k}}$
  - this works, despite the fact that we actually have different parent distribution functions $p(\mathbf{f}_{\mathbf{k}})$ for the different $\mathbf{k}$, since we can think of the sum as being a double sum: a sum over volumes in $\mathbf{k}$-space of size $\Delta k \ll 1/\Delta q$ – within which we have identical $p(\mathbf{f}_{\mathbf{k}})$ but still a very large number of modes – so we get a Gaussian distribution for each volume and we can then convolve these Gaussian distributions – which each have Gaussian generating functions, which get multiplied together to get a final Gaussian generating function – to get the above result.
  - finally, writing $f_{\mathbf{k}i} = \tilde{f}_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{q}_i}$ and (using the mirror symmetry of $\tilde{f}_{\mathbf{k}}$) $f_{\mathbf{k}j} = \tilde{f}^{\star}_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{q}_j}$ we get
    * $\boxed{M_{ij} = \sum_{\mathbf{k}} \langle \tilde{f}_{\mathbf{k}} \tilde{f}^{\star}_{\mathbf{k}} \rangle e^{-i\mathbf{k}\cdot(\mathbf{q}_i - \mathbf{q}_j)}}$
    * which unsurprisingly says that the components of the covariance matrix are what we would have gotten from the WK-theorem.
- it may be worth re-emphasising that the central limit theorem need not always apply
  - one can have statistically homogeneous random fields (or 'processes') which are *not* Gaussian
  - for which one can define autocorrelation functions and their relation to the power spectrum as above
  - the WK-theorem being independent of whether Gaussianity is obeyed
- but to the extent that the Gaussian model *is* applicable, it is an extremely useful tool
- in scintillation and imaging studies, for example:
  - if we can assume that the phase error $\phi(\mathbf{q})$ is Gaussian with structure function $S_\phi(\mathbf{d})$

- then one can show that the complex field amplitude on the observer plane $\psi(\mathbf{x})$ – a highly non-linear, non-local and non-Gaussian function of $\phi(\mathbf{q})$ – has 2-point function (the 'visibility' observed by radio astronomers) given by $\langle\psi(\mathbf{x})\psi^\star(\mathbf{x}+\mathbf{d})\rangle = \exp(-S_\phi(\mathbf{d})/2)$
- and from this, in the strong scattering regime, where $\psi(\mathbf{x})$ itself may be assumed to be a Gaussian random field, we find:
  - the flux density has an exponential distribution $P(F)dF = (dF/\overline{F})\exp\big(-F/\overline{F}\big)$ and
  - the structure function of the flux density $S_F(\mathbf{d}) = \langle(F(\mathbf{x}) - F(\mathbf{x}+\mathbf{d}))^2\rangle = 1 + \exp(-S_\phi(\mathbf{d}))$

- so the machinery of Gaussian statistics provides very powerful links between properties of remote scattering material and observable quantities