L3 Astro - Section 2

Nick Kaiser

Contents

List of Figures

1 Radiation Emission

1.1 Maxwell's equations

Maxwell's equations in the SI (rationalised) system are:

$$\nabla \cdot \mathbf{E} = \rho/\epsilon_0 \qquad \nabla \cdot \mathbf{B} = 0$$
$$\nabla \times \mathbf{E} = -\dot{\mathbf{B}} \qquad \nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \dot{\mathbf{E}}/c^2 \tag{1}$$

We have already discussed how this admit wave-like solutions in vacuum ($\rho = 0$ and $\mathbf{j} = 0$) These propagate at the speed of light and are, in general, polarized.

The inhomogeneous equations (giving $\nabla \cdot \mathbf{E}$ and $\nabla \times \mathbf{B}$) describe how such waves are generated and how they are detected.

1.2 The electromagnetic spectrum

1.2.1 The different wave-bands

- gamma rays (> ~1 MeV)
- hard X-rays (10-1000 keV)
- soft X-rays (1-10 A)
- EUV (~100 A)
- UV (~1000 A)
- visible (4000-7000 A -- 400-700 nm)
- near IR (~1 micron)
- IR (10 microns)
- THz (~100 microns--3000 GHz)
- submillimeter (300 GHz - 700 GHz)
- millimeter (30 GHz - 300 GHz)
- microwave (3 GHz - 30 GHz)
- decimeter (300 MHz - 3 GHz) ("cable" TV/UHF band)
- meterwave (30 MHz - 300 MHz) (TV/FM/HF band)
- dekameter (3 MHz - 30 MHz) (Shortwave
- AM band (0.5 MHz - 1.7 MHz)

etc.

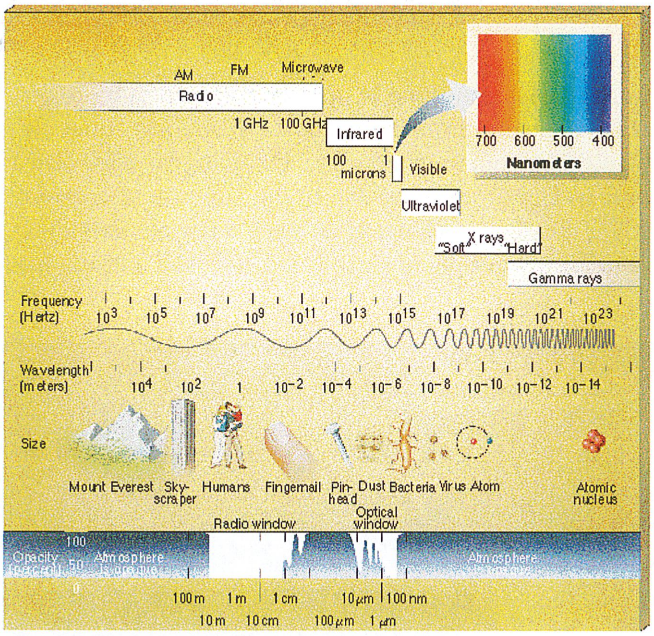

Figure 1: The electromagnetic spectrum and the terminology used to describe the different 'bands'.

1 eV ≃ 11000 K

$k_B T \simeq 1 eV$
$10^4 K$

1.2.2 The optical and radio atmospheric windows

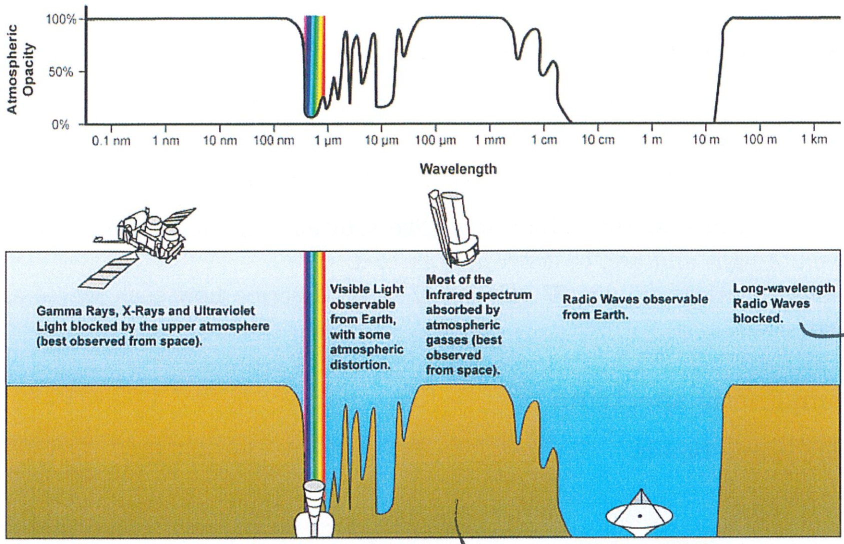

Figure 2: The atmospheric opacity is shown in more detail here: optical light occupies a rather narrow part of the spectrum around $\sim 4000\AA$. There are some other windows in the IR. Water in the atmosphere is a significant source of opacity. By going to high altitude, very dry, sites (e.g. $\sim$ 5000m in the Atacama desert in Chile) it is possible to do 'sub-mm' astronomy from the ground. The other important window is in the radio for wavelengths between a few cm and about 10m. Lower frequency radiation cannot propagate through the ionosphere since it is below the 'plasma frequency'.

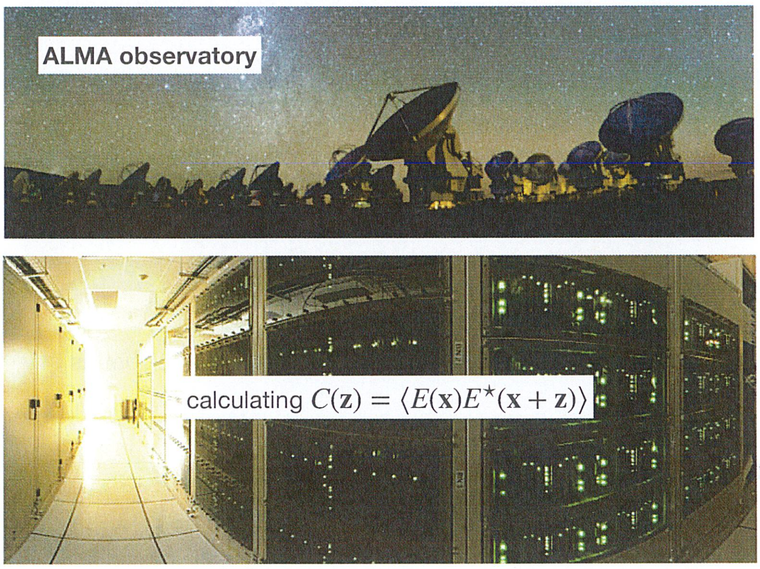

Figure 3: The ESO ALMA observatory. Situated in a very high (5km) and dry site in the Atacama desert in Chile, ALMA observes in the range $0.3\text{mm} \lesssim \lambda \lesssim 3\text{mm}$. The hardware is a combination of the 66 dishes, that sample the electric field coming from directions within the 'primary' beam with width $\theta \sim \lambda/D$, and the supercomputer below, which calculates the two point function of the electric field. The Fourier transform of this gives images with angular resolution $\theta \sim \lambda/L$ where L is the size of the array. It detects mostly emission from molecules, and probes deep into dense objects like molecular clouds.

1.3 Mechanisms for generation of EM radiation

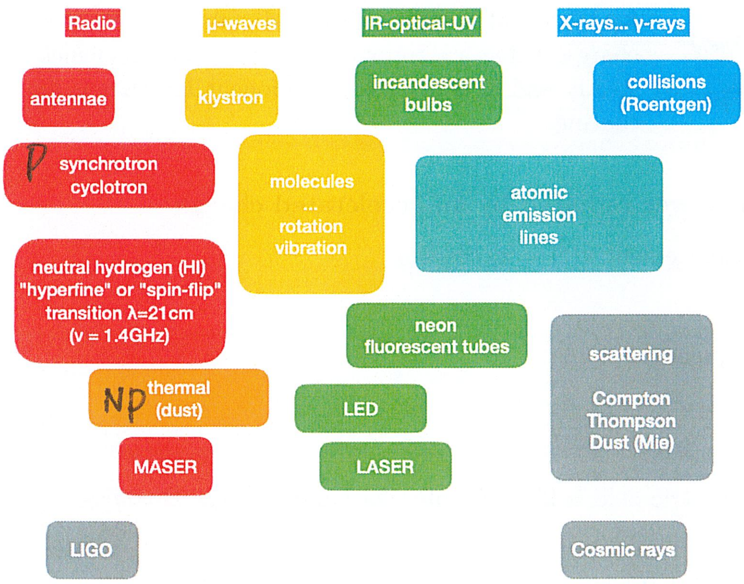

Figure 4: Sources of radiation used in astronomy (and elsewhere). Colour of boxes indicates the frequency. The grey boxes are not sources of EM radiation *per se* as scattering is re-processing of pre-existing radiation, and cosmic rays are particles, not EM radiation. Aside from the wavelength, an important qualitative distinction is whether the spectrum consists primarily of lines or 'continuum'. Another is whether or not the radiation is polarised. At bottom left is LIGO – the large interferometric gravitational wave observatory, whose $\nu \sim$ kHz observing frequency would put it way off to the left on this plot.

- radio – μ-waves – IR, visual, UV – X-ray – γ-rays
- artificial sources – natural sources – scattering
- what are the characteristics of spectrum – continuum or lines? – polarisation?

1.4 Radiation from a compact source

- How does the energy density vary at large distance ($\gg \lambda$ or source size)?
  - the energy flux density must fall off as $1/r^2$ (conservation of energy)
  - energy density is proportional to energy flux density ($v = c$).
- How does the *field strength* vary with distance?
  - $u = (|\mathbf{E}|^2 + |\mathbf{B}|^2)/2$ so $|\mathbf{E}|, |\mathbf{B}| \propto 1/r$
- how does this compare to electro- and magneto-static fields?
  - for monopole charge? or for a dipole (e.g. magnet)?
  - a compact charge distribution can be decomposed into monopole, dipole, quadrupole ...
    * monopole: $E \sim q/r^2$
    * dipole charge $E \sim qd/r^3$ (same r-dependence for dipole magnet)
    * quadrupole $E \sim qQ/r^4$
    * ...
  - all of these fall off much faster than the radiation fields

1.5 Radiation from an accelerated charge (Larmor's formula)

There are a wide range of processes that generate EM radiation

- What is the common feature of all of these processes?
- Does a charge in uniform motion radiate? Would this be compatible with special relativity?
- This leads us to suspect that it is *acceleration* of charges that gives rise to radiation.

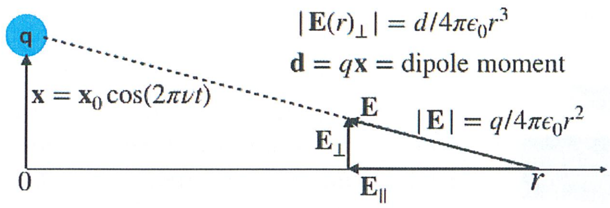

Figure 5: In the 'near field' – i.e. $r \ll \lambda = c/\nu$ of an oscillating charge with displacement $\mathbf{x}(t)$ the electric field is equal to that one would compute for a static charge. It has a component $\mathbf{E}_\parallel$ along the axis which is steady and a 'transverse' component $\mathbf{E}_\perp$ that is fluctuating.

1.5.1 Order of magnitude estimate for power radiated by an accelerated charge

- Consider an oscillating charge q with displacement $\mathbf{x}(t) = \mathbf{x}_0 \cos(\omega t)$
  - This is like a radio antenna
  - known empirically to radiate at frequency ω
  - or $\nu = \omega/2\pi$ in Hertz
  - it has a *dipole moment* $\mathbf{d} = q\mathbf{x}$
- At $r \ll \lambda$ – i.e. in the *'near-field'* – the electric field is like that of a stationary charge $\mathbf{E} = q\hat{\mathbf{r}}/4\pi\epsilon_0 r^2$ - approximately radial
- but it has a small *transverse* component with $|\mathbf{E}_\perp| = (x/r) \times (q/4\pi\epsilon_0 r^2)$ (see figure 5)
- so the transverse component of the field at the transition $r \sim \lambda$ is, to order of magnitude,
  - $E_\perp(r = \lambda) \sim d/4\pi\epsilon_0\lambda^3$
- at $r \gg \lambda$ – the *'radiation zone'* – we expect a transverse electric field $E_\perp(r)$ that falls of inversely with r (conservation of energy)
- *key assumption:* at the transition region where $r \sim \lambda$ we can use *either* the near-field or radiation zone formulae, so, for $r \gg \lambda$
  - $E_\perp(r) \sim (d/4\pi\epsilon_0\lambda^3) \times (\lambda/r)$
- if we square this we get the energy density $u \sim \epsilon_0 E_\perp^2$
- and if we multiply that by c we get the energy flux density $cu \sim \epsilon_0 c E_\perp^2$
  - which falls off as $1/r^2$ as befits an energy conserving flux density
- and if we multiply that by $4\pi r^2$, the area of a sphere at distance r, we get the total power radiated
  - $\boxed{P = cu \times (4\pi r^2) \sim x^2 q^2 c/\epsilon_0 \lambda^4 \sim \ddot{d}^2/\epsilon_0 c^3}$
- Key features:
  - power proportional to dipole moment (squared) and 4th power of frequency
  - or (charge × acceleration)^2

1.5.2 Larmor's formula from retarded potentials

- To obtain an accurate expression – including dimensionless factors of order unity – we need to use the *'retarded potentials'* of Lienard and Wiechert.
  - these are the exact solutions of Maxwell's equations for the electric and magnetic potentials $(\varphi(\mathbf{r}, t), \mathbf{A}(\mathbf{r}, t))$ of a moving charge or charge/current distribution $(\rho, \mathbf{j})$ where $\mathbf{j}$ is ρ times the mean velocity $\mathbf{v}$
  - the fields are $\mathbf{B} = \nabla \times \mathbf{A}$ and $\mathbf{E} = -\nabla\varphi - \dot{\mathbf{A}}$
  - they can be obtained by considering the potential to be the sum of the potentials generated by a set of infinitesimal spatial cells

- the cell at the spatial origin has contains a charge $q(t)$ which has a velocity $\mathbf{v}(t)$. It turns out that the potentials at position $\mathbf{r}$ and time t are the same as for a static charge and current element, but as they were at time $t' = t - r/c$; the retarded time
- so the potential – and hence the fields – 'here and now' are determined by the charge and current distribution on our 'past light cone'
- summing over cells gives
  - $\varphi(\mathbf{r}, t) = \frac{1}{4\pi\epsilon_0} \int d^3r' \rho(\mathbf{r}', t')/|\mathbf{r} - \mathbf{r}'|$
  - $\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0}{4\pi} \int d^3r' \mathbf{j}(\mathbf{r}', t')/|\mathbf{r} - \mathbf{r}'|$
  - where $t' = t - |\mathbf{r} - \mathbf{r}'|/c$

- if we consider a charge confined to a distance from the origin $|\mathbf{r}'|$ much less than the wavelength, then, to leading order, we can put $\mathbf{j}(\mathbf{r}', t') \simeq \mathbf{j}(\mathbf{r}', t - r/c)$
- and considering a point charge, so $\mathbf{j}(\mathbf{r}', t) = q\mathbf{v}(t)\delta^{(3)}(\mathbf{r}' - \mathbf{r}'(t))$
- we have $\mathbf{A}(r, t) \simeq (4\pi)^{-1}\mu_0 q\mathbf{v}(t - r/c)/|\mathbf{r} - \mathbf{r}'(t - r/c)| \simeq (4\pi)^{-1}\mu_0 q\mathbf{v}(t - r/c)/r$
- the $\mathbf{B}$-field is the curl (i.e. combination of spatial derivatives) of $\mathbf{A}$ which, for $r \gg \lambda$ (i.e. in the radiation zone), is dominated by the r-dependence of $\mathbf{v}(t - r/c)$
  - $B \sim \partial_r A \sim \mu_0 q \partial_r (v(t - r/c)/r) \sim q\dot{v}/cr = \ddot{d}/cr$
- and squaring this gives the energy density and we recover the same result as above
- putting in the factors of order unity results in *Larmor's formula:*
  - $\boxed{P = q^2 a^2 / 6\pi\epsilon_0 c^3}$
  - where a is the acceleration

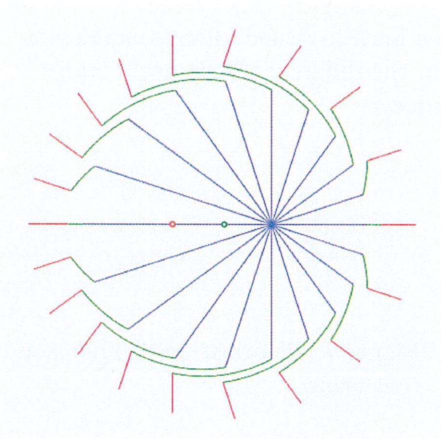

Figure 6: A commonly used figure to visualise the field from a moving charge. I think it is better to think of the magnetic potential from a small oscillating dipole. This has an overall radial fall-off with $A \sim \mathbf{j}/r$, but it has a fluctuating sign. The pattern of fluctuations moves out at the speed of light. The curl of $\mathbf{A}$ gives the magnetic field is on the order of $\mathbf{B} \sim \mathbf{A}/\lambda$, and also has an overall $\mathbf{B} \propto 1/r$ fall-off.

1.5.3 Implications of Larmor's formula

Larmor's formula has many applications and implications:

– **Radiation by (classical) atoms in Rutherford's model**

- if Larmor's formula were applicable to atoms – considered as being like little 'solar systems' with electrons orbiting the nucleus – the electrons would rapidly lose energy and spiral in to the nucleus
- this would be catastrophic for the theory

– **Scattering of light by molecules, dust grains ...**

This explains, in a general way, why is the sky blue?

- for a scattering charge which is 'tethered', the displacement (and therefore the dipole moment d) is proportional to the electric field E of the incoming radiation
- we can write the power in the scattered radiation as the incoming energy flux density (which is $\propto E^2$) times an area:
  - this defines the *scattering cross-section* σ
  - and more generally we can talk about a *differential* scattering cross section $d\sigma/d\Omega$ – a function of direction $\hat{\mathbf{n}}$ – giving the energy scattered per solid angle
- and, since $\ddot{d}^2 \sim \omega^4 d^2$, the total cross-section scales as the 4th power of frequency
- so higher frequencies are much more efficiently scattered

– **Scattering of light by electrons**

- called 'Thomson scattering' (if photon energy $\ll$ electron rest mass)
- here the electrons are not tethered, so the dipole is not $d \propto E$ but $d \propto E/\omega^2$
- so the cross section is *independent of frequency*
  - $\boxed{\sigma_{\mathrm{T}} = (6\pi)^{-1}(q^2/\epsilon_0 m_e c^2)^2 = 6.69 \times 10^{-29}\mathrm{m}^2}$
- unpolarised radiation scattered by electrons becomes polarised (see figure 7)
  - the degree of polarisation being proportional to the quadrupole moment of the incident radiation
  - so observing polarisation is a kind of *remote sensing* of the intensity incident on the scatterers
  - this provides an important diagnostic of the 'epoch of reionisation' when the first astrophysical sources ionised the previously neutral gas
- and polarised radiation has its polarisation modified
  - this being described by an equation of radiative transfer with a matrix-valued, directional dependent, cross-section that tells us how the energy flux density in the different polarisation states – i.e. the Stokes parameters – get modified in the scattering process

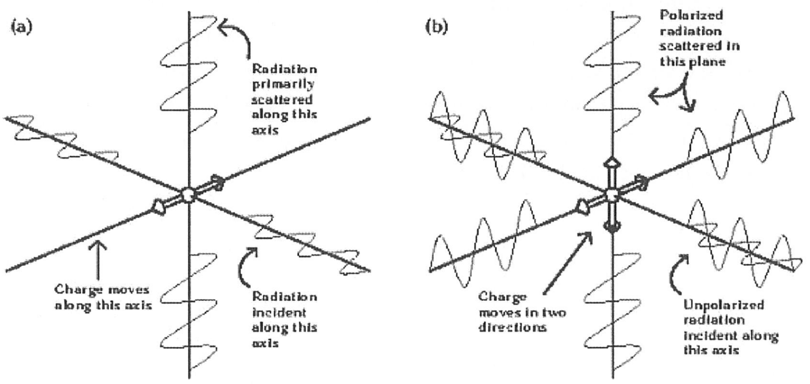

Figure 7: Polarisation of Thomson scattering.

1.5.4 X-ray emission from hot gas

- we will also use Larmor's formula when we compute compute the power radiated by electrons being deflected by ions in a hot plasma – so called *'thermal bremsstrahlung'*
  - the motion here is not simple harmonic oscillatory motion
  - the acceleration is a 'pulse' with width in time $\tau \sim b/v$ where b is the 'impact parameter'
  - and this should really be computed quantum mechanically
  - but we can still use Larmor's formula to compute the mean power radiated per collision and its distribution over frequency

1.5.5 Cyclotron and synchrotron emission

The mechanism by which electrons create radio emission is sketched in figure 8.

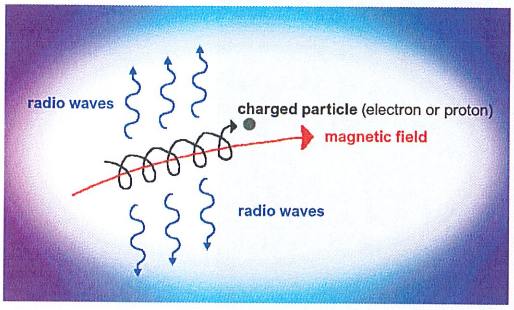

Figure 8: Electrons gyrating in a magnetic field are accelerated and therefore radiate according to Larmor. If they are non-relativistic this is called *cyclotron radiation* and if they are relativistic it is called *synchrotron radiation*.

- the equation of motion for a charge in a magnetic field is
- $\dot{\mathbf{p}} = q\dot{\mathbf{x}} \times \mathbf{B}$
- where dot denotes d/dt
- the relativistic 3-momentum is $\mathbf{p} = d\mathbf{x}/d\tau$ where $\tau = t/\gamma$ is proper time and $\gamma = 1/\sqrt{1 - v^2/c^2}$
- so this gives the proper acceleration, for a relativistic electron,
- $d^2\mathbf{x}_\perp/d\tau^2 = a_{\rm proper} = \gamma q c B$
- and so the power radiated (in the rest-frame of the electron) is
- $P \propto \gamma^2 B^2$
- but the power is a *Lorentz invariant*
  - it is $\Delta E/\Delta t$ with both energy and time transforming like time-components of a 4-vector

The radiation in the observer's frame is tightly beamed. So we would not see a simply sinusoidal field from a single charge, rather we would see a series of pulses of frequency boosted by a factor γ relative to the orbital frequency as illustrated in figure 9.

The fact that $P \propto \gamma^2$ while the energy of the particles is $E \propto \gamma$ means that the fastest electrons (those which emit the highest frequency radiation) lose energy more rapidly than lower energy ones.

So by measuring the cut-off in the spectrum at high frequency we can determine roughly when the electrons were ejected from the source.

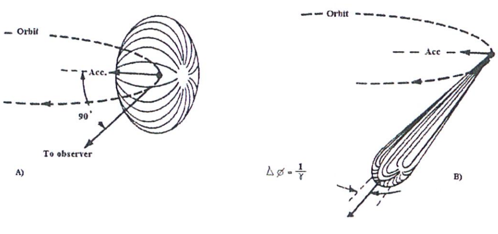

Figure 9: Accelerating charges radiate with a quadrupole pattern in their instantaneous rest-frame. In the observer's frame, and for a relativistic electron, this radiation becomes tightly 'beamed'

– **Some questions to consider:**

- Q: How would a molecule with no dipole moment, but with an oscillating *quadrupole* charge distribution radiate?
- Q: what about gravitational radiation? Can something like a black-hole binary have a dipole moment?

Figure 1. The giant radio galaxy Hercules A. Radio synchrotron jets emerging from the optical host of the galaxy mark the presence of magnetic fields roughly 1 million light-years in scale. Credit: NASA, ESA, S. Baum and C. O'Dea (RIT), R. Perley and W. Cotton (NRAO/AUI/NSF), and the Hubble Heritage Team (STScI/AURA)

Figure 10: Radio image of the 'double-lobed' synchrotron radiation emission from relativistic electrons emerging from the nucleus of the elliptical galaxy Hercules-A.

2 Radiation and absorption by atoms and molecules

2.1 Atomic physics before quantum mechanics

- Greeks ... Dalton ... Avagadro/Loschmidt
  - Dalton is famous for the atomic theory
    - ∗ what was the evidence for this?
  - Avagadro's number was determined in the early 19th century and gave the mass of an atom
    - ∗ how?
  - and the size of the atom was known to be about $\sim 10^{-10}$m ($1\AA$)
    - ∗ how did they know all this?
    - ∗ was it universally accepted?
- Electrons were known to be *particles*, not waves
  - Crooke's experiment – sharp shadow
  - J.J. Thomson: cathode rays carry charge - discharge of gold-leaf electroscope
  - and they have a high charge-to-mass ratio (easily deflected)
  - he proposed they live in atoms, with positive charge smoothly distribited
    - ∗ the 'plum-pudding' model
- Rutherford model of the atom
  - he scattered α-particles off a thin gold film
    - ∗ some of these were scattered to large angle
  - the positive charge is concentrated in a very small volume (the nucleus)
    - ∗ suggests analogy with gravitation in the solar system with inverse-square gravitational attraction replaced by inverse squared electrical attraction
  - Q: in a hydrogen atom, what accelerates more – the electron or the nucleus? If the electron and nucleus had the same mass, how would this affect the 'catastrophe'.

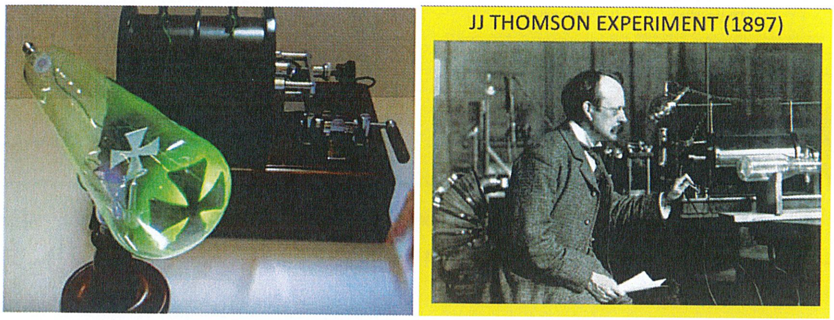

Figure 11: Crooke's experiment (left) showed that cathode rays cast sharp shadows. J.J. Thomson performed numerous experiments to show that cathode rays carry electric charge and thus must be particles. He got the Nobel proze for this. His son G.P. got the Nobel prize for showing that they are waves.

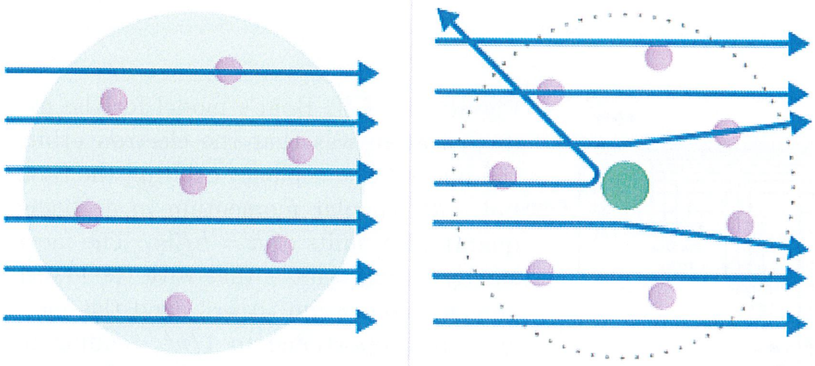

Figure 12: Following Crooke's and Thomson's experiments the prevailing model for atomic structure was the 'plum pudding' model in which electrons were particles distributed in a spread out positive charge distribution. Rutherford's gold-film α-particle scattering experiments showed that the positive charge had to be concentrated in a tiny nucleus.

2.2 Quantum mechanics & the 'Bohr-atom' model

Atoms were known to emit and absorb radiation at discrete frequencies, with patterns in frequency: $\nu_{nm} = \text{constant}(1/n^2 - 1/m^2)$. The first (though somewhat flawed) explanation of this was by Niels Bohr – the 'Bohr atom'.

2.2.1 The Bohr-atom

- for classical point electron particle orbiting a hydrogen nucleus:
  - $F = ma \quad \Rightarrow \quad e^2/4\pi\epsilon_0 r^2 = m_e\omega^2 r = m_e v^2/r$
- *If* angular momentum is *quantized* in units of $\hbar \quad \rightarrow \quad m_e vr = n\hbar$
  - multiply $F = m_e a$ by $m_e r^3$ to eliminate v:
  - $\Rightarrow \quad m_e e^2 r/4\pi\epsilon_0 = m_e^2 v^2 r^2 = n^2\hbar^2$
  - gives $\boxed{r_n = 4\pi n^2\hbar^2\epsilon_0/m_e e^2}$
  - so energy $\sim 1/n^2$
- Q: how big is a hydrogen atom?
  - A: $r_1 = 4\pi\hbar^2\epsilon_0/m_e e^2 = 0.53\text{Å}n^2$
  - Q: what about singly ionized helium? or other single electron ions with nuclear charge Ze?
- Q: what are the energy levels?

$$E = -\frac{1}{2}\frac{Ze^2}{4\pi\epsilon_0 r} = -\frac{1}{2} \times \frac{Ze^2}{4\pi\epsilon_0} \times \frac{m_e e^2 Z}{4\pi\epsilon_0\hbar^2} \times \frac{1}{n^2} = -\frac{1}{2}m_e c^2\frac{\alpha^2 Z^2}{n^2} \tag{2}$$

  - the factor $-1/2$ coming from the 'virial theorem': $2T + V = 0 \Rightarrow E = T + V = -V/2$
- where $\boxed{\alpha \equiv e^2/4\pi\epsilon_0\hbar c \simeq 1/137}$ is the *fine structure constant*
  - $\alpha \ll 1$ so energy $\ll$ electron rest-mass energy $m_e c^2$:

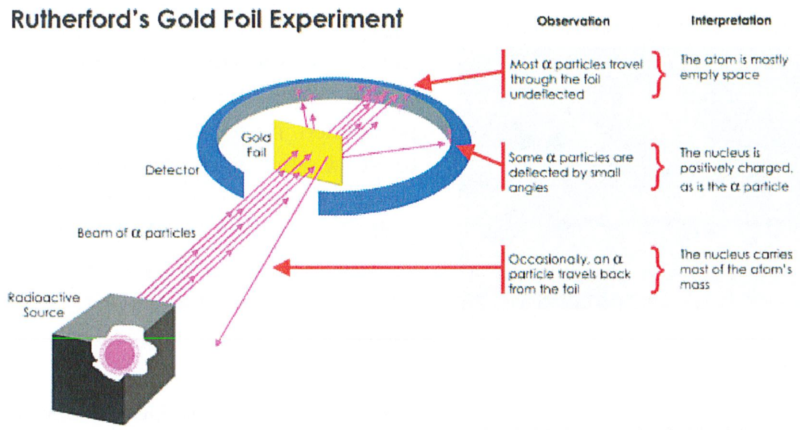

Figure 13: Rutherford measured scattering of α-particles from gold atoms.

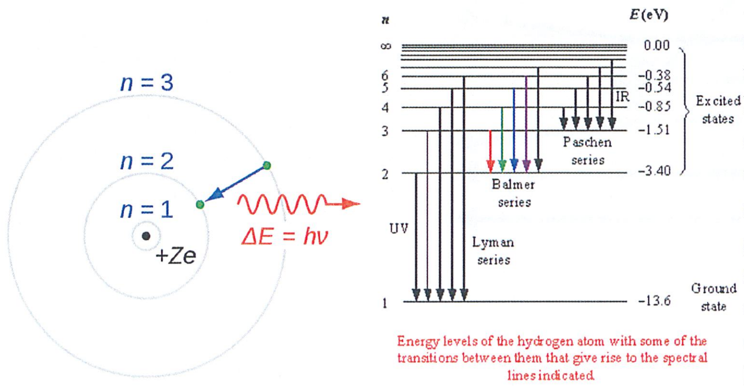

Figure 14: Niels Bohr's model for the hydrogen atom was that the electron orbits the nucleus like a planet orbiting the Sun, but with angular momentum $m_e v r$ being quantised in units of $\hbar = h/2\pi$. The most tightly bound orbit is that with the lowest angular momentum ($n = 1$) and the energies are proportional to $1/n^2$, tending to zero as $n \to \infty$. The energy of a photon emitted (or absorbed) is $h\nu = \Delta E$, the difference in energies between the orbits. This explained the empirical observation that $\nu = \text{constant} \times (1/n^2 - 1/m^2)$.

– so the non-relativistic Schroedinger equation is a good description (for low Z at least)

Features of the "Bohr-atom":

- wrong in detail (e.g. for the angular momentum) but gives correct energy levels
  - with $h\nu_{mn} = E_n - E_m$ explains patterns seen by spectroscopists
  - the frequency of light emitted (or absorbed) in transition $n \to m$

2.2.2 Spectroscopist's terminology

- *ground state* $n = 1$
- *Lyman series* $n \geq 2 \to n = 1$
  - Lyα: $n = 2 \to 1$; Lyβ: $n = 3 \to 1$; ...
  - *Lyman limit:* $n = \infty \to 1$
    * a photon emitted in (re)combination of ion plus a low-energy free electron to ground state has energy $\boxed{E = \alpha^2 m_e c^2/2 \simeq 13.6\text{eV}}$ and wavelength$\lambda = 912$Å.
    * photon energy needed to dissociate (ionize) a H-atom in the ground state
    * also called the *Rydberg energy* or the *ionisation potential*
  - Lyman transitions are in the ultra-violet band
- *Balmer series* $n \geq 3 \to n = 2$
  - usually called Hα ($n = 3 \to 2$, $\lambda = 6563$Å); Hβ ($n = 4 \to 2$, $\lambda = 4861$Å)
  - these transitions give/absorb optical photons (H_α gives the red glow of nebulae)

2.2.3 Radiation vs. orbital frequency

- note that the classical velocity $v_c = \sqrt{2E/m}$ $\rightarrow \hbar\omega_c = \hbar v_c/r = \alpha^2 m_e c^2/n^3$
  - the frequency of radiation radiated according to Larmor for a classical particle orbiting with angular frequency ω_c
  - approximates the Bohr frequency for transition $n \rightarrow n-1$ for $n \gg 1$
  - so a highly excited H-atom making transitions $n \rightarrow n-1 \rightarrow n-2 \ldots$ radiates quasi-classically

2.2.4 Radiation from atoms in the Schrödinger picture

The *probability current* times the charge is

$$\mathbf{j} = \frac{iq\hbar}{2}(\psi^\star \nabla \psi - \text{c.c.}) \tag{3}$$

- Energy eigenstates of Schrödinger equation have steady currents so, classically, would not be expected to radiate.
- But a *superposition* of two eigenstates $\psi = \psi_n + \psi_m$ *does* have fluctuating current, charge density.
- And it fluctuates at the frequency given by the energy difference.

A more accurate treatment would involve:

- (non-relativistic) Schrödinger equation: proper accounting for orbital angular momentum
- relativistic corrections
- electron/proton spin-flip transitions (radio frequency)
  - these give *hyperfine* transitions
  - and particularly the neutral hydrogen (HI) line at $\lambda = 21$cm
  - very important in radio astronomy

– **Calculation of rates of transitions (schematic)**

- we treat the modes of the radiation as simple harmonic oscillators
  - $H = m\dot{x}^2/2 + kx^2$
  - has energy eigenstates $|n\rangle$ with $E_n = (n + 1/2)\hbar\omega$
  - can be generated by applying the creation and destruction operators:
  - $a^\dagger, a \sim x \pm ip$ (in suitable units)
  - $|n+1\rangle = \sqrt{n+1}a^\dagger |n\rangle$
    $|n-1\rangle = \sqrt{n}a\,|n\rangle$
- Set up eigenstates for radiation (occupation number eigenstates $|\ldots n_\mathbf{k} \ldots\rangle$) and atom (orbital quantum number n)
- The classical interaction energy is $H_{\text{int}} = \int d^3x \mathbf{A} \cdot \mathbf{j}$. Treat this as a *perturbation* in Schrödinger's equation
  - i.e. replace $\mathbf{A}$ and $\mathbf{j}$ by operators
  - use the 'interaction picture' (hybrid of Heisenberg and Schrödinger pictures)
    * operators have time dependence of unperturbed eigenstates (like Heisenberg)
    * states evolve (like in Schrödinger) but only due to the perturbation
- Schrödinger equation:

- $$\boxed{i\hbar\frac{d}{dt}|\ \rangle = H_{\rm int}|\ \rangle}$$
  - solve this iteratively
- to zeroth order nothing happens: initial state $|i\rangle$ is unchaanged
- to first order $|i\rangle \to |t\rangle = |i\rangle + i\hbar \int dt\ H_{\rm int}\,|i\rangle$
- calculate - to 1st order- amplitude $\langle f|t\rangle$ to be in a different state $|f\rangle$
- i.e. amplitude for transition like $n \to m$ and $n_{\bf k} \to n_{\bf k} + 1$:
- $\langle m|\,\langle 1_{\bf k}|\, i\hbar \int dt H_{\rm int}\,|0_{\bf k}\rangle\,|n\rangle$
- this contains $\langle m|\,{\bf j}\,|m\rangle$ which oscillates at frequency $\hbar\omega = E_n - E_m$
- and gives non-vanishing amplitude if the radiation mode $\bf k$ has that frequency
- square and divide by time to get the *rate of transitions*
  - most rapid are *electric-dipole* transitions: where $\langle n|m\rangle$ has a fluctuating electric-dipole moment
  - but with very low density astrophysical plasmas we also observe *electric quadrupole* and *magnetic-dipole* transitions
  - not seen in the laboratory as collisions 'de-excite' more rapidly
  - for this reason these are called *'forbidden'* lines

3 Stellar atmospheres and spectra

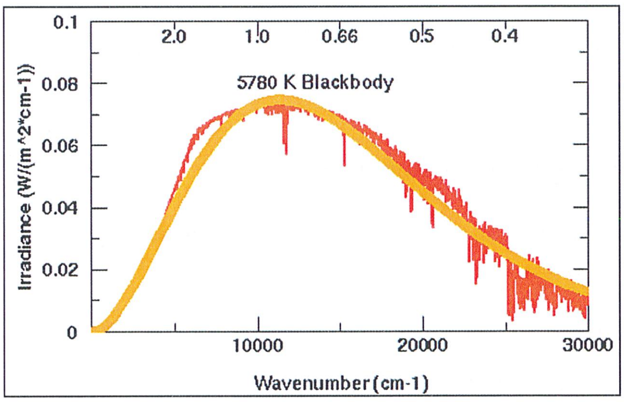

Figure 15: Spectrum of the sun compared to a black-body.

- to a crude – but often very useful – approximation stars emit black body radiation
  - characterized solely by the temperature T
  - T can be determined from a single 'colour'
  - for 'main-sequence' stars this (and the size) gives the intrinsic luminosity L
  - and hence distance from observed flux-density
- a good example is the sun (see figure 15)
- but spectra of some other types of stars have significant departures from BB form (see figure 16)
  - primarily absorption as radiation propagates through the stellar atmosphere
  - as seen in the solar spectrum at around 4000 $\mathring{A}$

- requires solution of *equation of radiative transfer*
  - *transport* equation for the brightness I_ν
  - in vacuum, I_ν is constant along a path
  - deep inside a star, matter is in thermal equilibrium at the same temperature as the radiation so $I_\nu = B_\nu$
  - but in the atmosphere – where *optical depth* is small or modest – absorption/emission by atoms and *scattering* modify the emitted spectrum
- Practical implications:
  - allows determination of the chemical composition of the (atmosphere) of the star
  - discrimination between MS and non-MS stars from spectra
    - e.g. by determination of the surface gravity via pressure broadening
  - determination of *photometric redshifts* for galaxies from e.g. the 'H & K break'
    - Calcium ions ⇒ H and K-lines
    - observationally inexpensive – obtained from broad-band colours

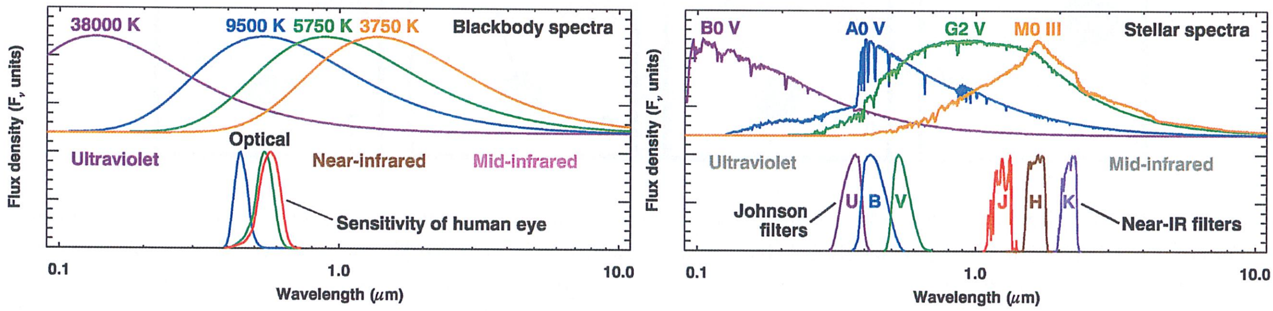

Figure 16: On the left are shown some black-body (i.e. thermal) spectra and an indication of the wavelengths of light that the human eye is sensitive to. On the right are shown some spectra of actual stars – some of which differ substantially from black-body form. On the bottom are shown the 'transfer functions' giving the transmission of filters used by astronomers.

Atomic emission is also important in nebulae. E.g. the prominent red glow of H-α. Note that excitations can be *radiative* or *collisional*.

4 Saha's equation for equilibrium ionization & excitation for a plasma

4.1 Inferring the composition of stellar atmospheres: the problem

- Stars exhibit absorption features that derive from hydrogen and other atoms.
- From such measurements one would like to infer the chemical composition of stars.
- in order to do this, one needs to know, in addition to the absorption cross-section
  - given the temperature, what fraction of the atoms of each element will be neutral?
    - a pre-requisite for absorption
  - and if we are dealing with transitions from an excited level, we need to know. what fraction of the neutral atoms are in that excited state?
- Lets consider hydrogen for simplicity: What is the physical state of hydrogen in the atmosphere of a star like the sun say. There are two questions:
  1. what is the fractional excitation? e.g. $f(n=2)/f(n=1)$ $\rightarrow$ Balmer lines
  2. what is the fractional ionization? (or equivalently neutral fraction $f_{\rm neutral} = 1 - f_{\rm ionised}$)

- the answer to (1) given by Boltzmann's formula:
  - for example, the relative abundance of hydrogen atoms in the first excited state ($n = 2$) to the ground state ($n = 1$) is $n_2/n_1 = \exp(-\Delta E/k_B T)$.
  - One electron volt corresponds to a temperature of $\simeq 11,000$K
  - $\Delta E(2-1) \simeq 10$eV $\sim 10^5$ K.
    * Q: what is excitation of the Sun ($T \simeq 6000$K)?
- (2) is more difficult
  - naively one might use Boltzmann's formula with $\Delta E = \chi \simeq 13.6$eV
    * would predict very low ionization
  - another approach might be to set up a system of differential equations with rates for ionization, recombination etc
    * but this would be very arduous

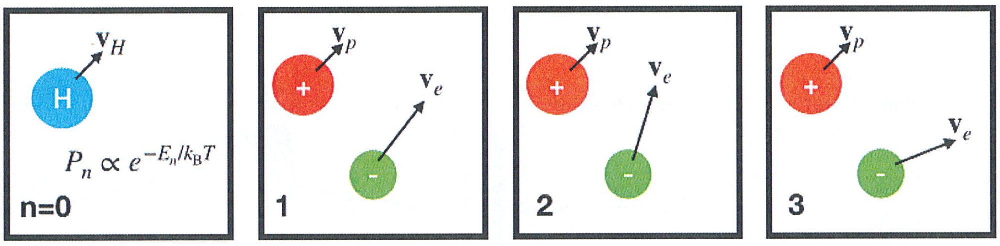

Figure 17: Illustration of Saha's calculation. The idea is that a 'sub-system' of a larger subsystem will occur, in thermal equilibrium, with probability given by Boltzmann's law. Here the sub-system is one proton and one electron. This may be found in the neutral state (a Hydrogen atom) as shown in the left panel (n=0). Or it may be found to be ionized, as in the other panels. The probability for any of these is smaller than the probability of the neutral state, since they have higher energy, but to obtain the net probability to be ionized we need to sum over all possible ways for the sub-system to be ionized (with weights proportional to $\exp(-E/k_B T)$). The result, as Saha showed in 1920, is that the ionisation fraction is greatly enhanced with respect to a naive application of Boltzmann's law.

4.2 Saha's solution

- assume thermal equilibrium: use Boltzmann's law to calculate the relative probability to find a proton + electron to be in either one of the neutral, but possibly excited, states or in an ionized state
  - but allow for the fact that there are, crudely speaking, many ways more ways for a proton-electron pair to be ionized as there are many possible values for the momentum of the electron
  - the upshot of this is that hydrogen becomes ionized at a much lower temperature than that at which $k_B T$ is equal to the ionization potential
- consider a volume containing, on average, one free electron
  - $V = L^3 = 1/n_e$
  - so $\Delta k = 2\pi/L \quad \Rightarrow \quad \Delta p = \hbar \Delta k = h/L \quad \Rightarrow \quad \Delta v = h/lm$
- Boltzmann: $n(\mathbf{v})/n_1 = \exp\big(-(\chi + mv^2/2)/k_B T\big)$
  - where n_1 is the number density of atoms in the ground state
  - and $\chi = 13.6$eV is the ionization potential
- but there are many possible values for $\mathbf{v}$, so the net probability is higher
- Q: how many possible $\mathbf{v}$ values are there for the electron?

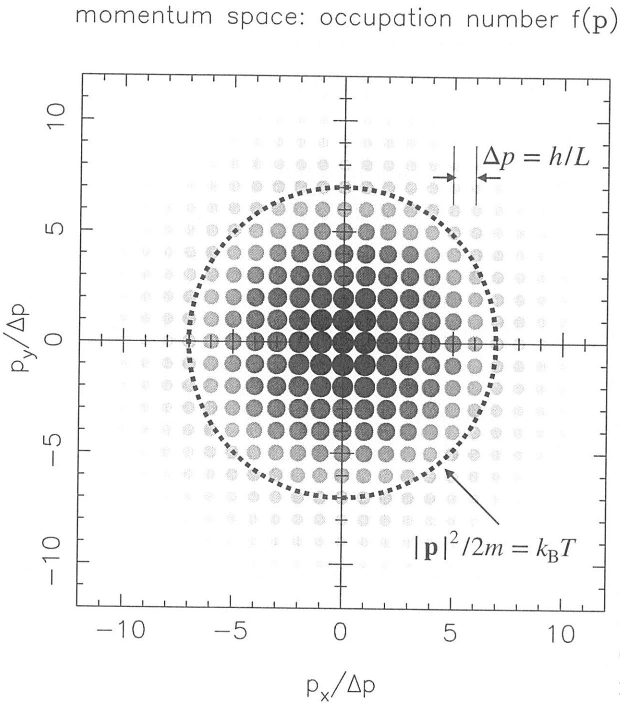

Figure 18: A non-degenerate gas of particles in thermal equilibrium at temperature T has mean occupation number $f(\mathbf{p}) \propto \exp(-E(\mathbf{p})k_\mathrm{B}T)$ where $E(\mathbf{p}) = |\mathbf{p}|^2/2m$; i.e. a *Maxwellian* or *Maxwell-Boltzmann* distribution. The de Broglie waves for the particles are assumed to be periodic inside a large box of side L, the spacing of the allowed modes in *wavenumber space* $\mathbf{k}$ – just like the allowed standing waves for radiation in a box with reflecting walls – is $\Delta k = 2\pi/L$ so the spacing of modes in *momentum space* is $\Delta p = \hbar\Delta k = h/L$. The plot shows the allowed modes in the plane $p_z = 0$. The area (and darkness) of the circles is proportional to f. The dashed circle indicates the value of the momentum such that the kinetic energy of a particle is equal to $k_\mathrm{B}T$. The Maxwell-Boltzmann distribution function $f(\mathbf{p})$ is that which maximises the *entropy* $S = \log W$ where $W = \prod_s A^s!/\prod_r p_r^s!$, where A^s is the number of modes on the s^{th} shell. Much as Bose calculated for light quanta – but where non-degeneracy implies most cells are empty so in the denominator we only consider $r = 0$ or $r = 1$ and we obtain $S = -\sum_s A^s f_s \log f_s$ where $f_s = p_1^s/A^s$.

- classically, the answer would be infinite
- but quantum mechanically the allowed *de Broglie wave modes* $\psi \sim \cos(\mathbf{k}\cdot\mathbf{x})$ live on a cubical lattice in $\mathbf{k}$ space with spacing $\Delta k = 2\pi/L$ (see figure 18)
- a lot like we found for the allowed states for the radiation field (standing waves)
  * this is a very common trick throughout astrophysics (and more generally)
  * we assume that the fields in the universe are periodic within some large box of side L
  * this is for computational convenience (the value of L drops out of the final results – or should!)
- this gives a lattice of momentum states with spacing $\Delta p = \hbar\Delta k = h/L$.
- and so the allowed velocities live on a lattice with spacing $\Delta v = \Delta p/m$

• so to get n_+/n_1 (where n_+ is the number density of ions) we need to sum over the allowed discrete velocity states (de Broglie wave modes) weighted by the Boltzmann probability
  - $n_+/n_1 = \sum_{\mathbf{v}} n(\mathbf{v})/n_1 = \sum_{\mathbf{v}} \exp(-(\chi + mv^2/2)/k_\mathrm{B}T)$
  - splitting this exponential into two factors and judicuously introducing factors of $(\Delta v)^3$ we have
    * $n_+/n_1 = \exp(-\chi/k_\mathrm{B}T)/(\Delta v)^3 \times \sum(\Delta v)^3 \exp(-mv^2/2k_\mathrm{B}T)$
  - where the sum can now be replaced by an integral ($\sum(\Delta v)^3 \ldots \Rightarrow \int d^3v \ldots$)to obtain
    * $n_+/n_1 = \exp(-\chi/k_\mathrm{B}T)/(\Delta v)^3 \int d^3v \exp(-mv^2/2k_\mathrm{B}T)$
  - but $\int d^3v \exp(-mv^2/2k_\mathrm{B}T) = 4\pi(k_\mathrm{B}T/m)^{3/2} \int dy\, y^2 e^{-y^2/2} \sim (k_\mathrm{B}T/m)^{3/2}$
    * since the dimensionless integral here is just a number of order unity

• and using $\Delta v = h/mL$ we have $n_+/n_1 \sim (Lm/h)^3(k_\mathrm{B}T/m)^{3/2}\exp(-\chi/k_\mathrm{B}T)$

• and finally, since since we chose the volume to be that containing on average 1 electron, we have $L^3 = 1/n_e$ and hence we have *Saha's equation* for the fractional ionisation
  - $\boxed{n_+/n_1 \sim (\lambda_{\mathrm{dB}}^3 n_e)^{-1}\exp(-\chi/k_\mathrm{B}T)}$
  - where $\lambda_{\mathrm{dB}} = h/p(T) = h/\sqrt{m_e E(T)} = h/\sqrt{m_e k_\mathrm{B}T}$ is the typical de Broglie wavelength of a thermal electron (KE $\sim k_\mathrm{B}T$)

- Comments and features of Saha's equation:
  - a pioneering application of quantum mechanics and statistical mechanics to astrophysics
  - the dimensionless pre-factor $1/\lambda_{\rm dB}^3 n_e$ gives a *much* higher ionization than the naive expectation
    - for low-density (highly non-degenerate) plasmas
  - note that what we have calculated is ratio of ionized (n_+) to *ground state* (n_1)
    - to get the actual $n_+/n_{\rm neutral}$ we would need to add the other excited states in $n_{\rm neutral} = n_1 + n_2 + \ldots$
    - this is fairly straightforward
    - and for e.g. the Sun it has little impact as most of the neutral atoms are in the ground state
  - and we have only considered hydrogen (for simplicity, but it is an important case)
    - for higher atomic number, a straightforward generalisation gives the fraction of atoms in the various ionisation states (singly ionised, doubly ionised ...)

4.3 Implications of Saha's equation

- The observed absorption in the Sun implies a much greater hydrogen abundance than was previously thought
  - hydrogen is the dominant component of the Universe
- It explains why Balmer absorption lines are most prominent in stars with $T \sim 10^4$K (see figure 19)
  - for higher T the neutral fraction drops precipitously
  - for lower T the fraction of atoms that are in 1st excited state is negligible
- Another important implication of Saha's equation is in cosmology:
  - In the early universe the temperature was very high and all of the matter was highly ionized and the radiation was tightly coupled to the matter
  - but as the universe cooled, the primordial plasma 'recombined' and became a gas of neutral matter
  - at that time the photons of the cosmic microwave background (CMB) 'de-coupled' and propagated essentially freely to the present epoch
  - the CMB radiation provides a 'snapshot' of the universe at the epoch of recombination
  - to interpret that we need to know when recombination occurred
  - naively one might estimate that to be the time when the thermal energy $k_{\rm B}T$ was equal, to order of magnitude, the ionisation potential $E \simeq 13.6$eV. Or about 10^5 K.
  - but just as with the Sun a naive Boltzmannian estimate of the ionisation fraction on the surface of the Sun that would give the wrong answer: in reality the universe remains ionised to a much lower temperature (about 3000 K).

A Statistical mechanics of non-degenerates gases and plasmas

Here we derive some useful expressions for the phase-space density (i.e. the mean occupation number) and the entropy for a non-degenerate thermal gas. We introduce the concept of density of states in momentum-space. We write down the 'complexion' *à la* Bose for a gas of indistinguishable particles and we express its logarithm – which is the statistical mechanical entropy – in terms of the mean occupation number $f(p)$ (really the phase space distribution function). We then show that this is maximised – subject to the constraints of a given number of particles and total amount of energy – for the 'Maxwell-Boltzmann' distribution $f \propto \exp(-E(p)/k_{\rm B}T)$ with mean energy per particles being $(3/2)k_{\rm B}T$.

A key result here is that the entropy per particle is minus the mean of the log of the occupation number. We will consider degenerate gases and plasmas in the next lecture.

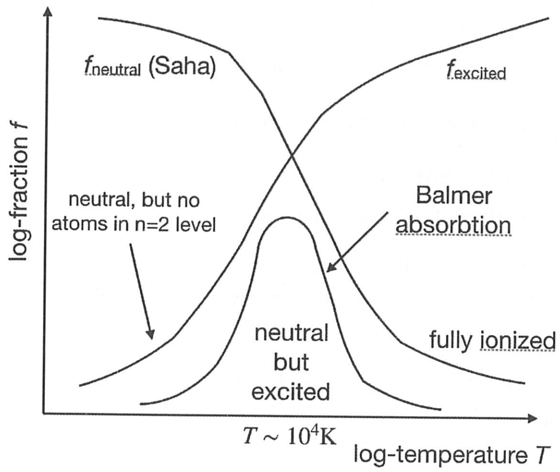

Figure 19: Quantities relevant to calculate the strength of Balmer absorption lines from the hydrogen in the atmosphere of a star as a function of the surface temperature T. The curve labelled $f_{\rm neutral}$ is the fraction of the hydrogen that is 'neutral' (i.e. non ionized). Only neutral hydrogen gives rise to absorption lines. At sufficiently high temperatures – though lower than the $\sim 10^5$ K one would infer from a naive application of Boltzmann's formula – hydrogen will be fully ionised and one will see no absorption. The curve labelled $f_{\rm excited}$ indicates what fraction of the non-ionised atoms are in the first excited state $n = 2$, which is an increasing function of temperature. The fraction of hydrogen that can absorb at Balmer frequencies is the product of these curves; it has a peak at around 10,000 K - somewhat hotter than the Sun.

A.1 Mean occupation number: single species of particles

A.1.1 Density of states

- We consider a universe within a large cubical volume $V = L^3$ with 'periodic boundary conditions'
  - the volume drops out in final results for things like density of particles, energy, entropy etc.
- so the allowed possible values for the wave-number $\mathbf{k}$ of a 'de Broglie wave' describing a particle lie on a lattice in $\mathbf{k}$-space with spacing $\Delta k = 2\pi/L$ (see figure 18)
  - so the possible values of the momentum $\mathbf{p} = \hbar\mathbf{k}$ lie on a lattice with spacing $\Delta p = h/L$
  - so we have one momentum state per phase-space volume $d^3p d^3x = \hbar^3$
  - this is just as for EM waves in a cavity
    - the fact we are dealing with massive particles rather than massless photons has no impact
    - though we may have different numbers of 'spin-states' per mode
    - but that just introduces a multiplicative factor for non-degenerate particles
  - 'non-degenerate' here means that we are considering situations where the number of particles per mode – or per volume $\hbar^3$ of 6-dimensional phase-space – is very small

A.1.2 The complexion and the entropy

- And, again as for radiation quanta, and following Bose's notation, we can define a 'complexion'
  - $W = \prod\limits_s W^s = \prod\limits_s (A^s!/\prod_r p_r^s!)$
  - where the integer s labels the shell and is related to the momentum $|\mathbf{p}| = s\Delta p$
  - A^s is the number of modes (or 'cells') in the $s^{\rm th}$ shell
    - $A^s = 4\pi g s^2$
    - where g is the number of spin-states per momentum state (e.g. $g = 2$ for electrons)
  - and p_r^s is the number of cells in the $s^{\rm th}$ containing r particles
  - so W^s, is the number of different sets of occupation numbers $\{r_j\}$, where j labels the cells, having a certain distribution of occupation numbers p_r^s
- and define the statistical mechanical entropy to be $S = \log W$ and, assuming that A^s and p_r^s are large numbers, we can invoke Stirling's theorem to obtain
  - $S = \sum\limits_s S_s = \sum\limits_s (A^s \log A^s - \sum\limits_r p_r^s \log p_r^s)$

- note that at any value of the (non-relativistic) energy $E(p) = p^2/2m = (h^2/mL^2)s^2/2$ (or of the momentum $p = hs/L$) the numbers s, and hence A^s and p_r^s, increase as we increase the size of the box, so in the limit $L \to \infty$ Stirling's approximation becomes arbitrarily accurate

- The differences, for particles such as atoms, or ions and electrons in a plasma are as follows:
  - if the particles are fermions, the occupation number r can only be 0 or 1
    - * though as for radiation quanta we may have multiple spin-states per cell
  - and for a highly non-degenerate gas (be it composed of fermions or bosons – like He^4) the probability that a cell contain a single particle is $P_s(r=1) = p_1^s/A^s$ is very small, and the probability to have $r > 1$ (zero for fermions) is negligible.

- with either of these restrictions, and defining the mean occupation number for the $s^{\rm th}$ shell to be $f_s \equiv p_1^s/A^s$ (with complement $1 - f_s = p_0^s/A^s$) the entropy is
  - $S = -\sum\limits_s A^s \sum\limits_{r=0,1} (p_r^s/A^s)\log(p_r^s/A^s)$
  - or
  - $\boxed{S = -\sum\limits_s A^s[(1-f_s)\log(1-f_s) + f_s \log f_s]}$
  - which is interestingly similar to the entropy for radiation (bosonic, rather than fermionic particles):
  - $S = \sum\limits_s A^s[(1+f_s)\log(1+f_s) - f_s \log f_s]$

- or, for sufficiently low occupation number (so $|\log f_s|$ becomes very large)
  - $\boxed{S = -\sum\limits_s A^s f_s \log f_s}$

- note that since the total number of particles is $N = \sum\limits_s A^s f_s$ this says that the entropy per particle is
  - $S/N = \langle -\log f_s \rangle$
  - where $\langle \ldots \rangle$ denotes the particle number weighted average

- or, in more physical terms, since the number of cells in a shell at momentum $p = |\mathbf{p}|$ and width $d\mathbf{p} = \Delta p ds$ is $A^s ds = 4\pi g p^2 dp/(\Delta p)^3 = 4\pi g L^3 h^{-3} p^2 dp$,
  - $\boxed{S = -4g\pi L^3 h^{-3} \int dp p^2 f(p) \log f(p)}$

- where $f(p = hs/L) = f_s$

A.1.3 The thermal distribution function

- The thermal distribution function f_s is that which maximises the entropy $S = -\sum_s A^s f_s \log f_s$ subject to the constraints on the total number of particles:
  - $N = \sum_s A^s f_s$

- and of the total energy
  - $E_{\rm tot} = (h^2/2mL^2)\sum_s A^s f_s s^2$

- introducing Lagrange multipliers α and β, the desired distribution must satisfy
  - $\delta(S - \alpha N - \beta E_{\rm tot})/\delta f_s = A^s[-(1 + \log f_s) - \alpha - \beta h^2 s^2/2mL^2] = 0$

- with solution
  - $f_s = e^{-(1+\alpha+\beta h^2 s^2/2mL^2)} = e^{-(1+\alpha)} e^{-\beta p^2/2m}$

- which, since $p^2/2m = E$, is evidently a Boltzmann distribution
- we can eliminate the normalisation factor $e^{-(1+\alpha)}$ by making the substitution $\sum_s A^s \ldots \to 4\pi(L/h)^3 \int dpp^2 \ldots$ in $N = \sum_s A^s f_s$:
  - $N = 4\pi L^3 h^{-3} \int dpp^2 f_s = 4\pi L^3 h^{-3} e^{-(1+\alpha)} \int dpp^2 e^{-\beta p^2/2m} = 4\pi(L/h)^3 e^{-(1+\alpha)} (m/\beta)^{3/2} \int dyy^2 e^{-y^2/2}$
- or, with $\int dyy^2 e^{-y^2/2} = \sqrt{\pi/2}$ and $N = nL^3$, with n the number density of particles,
  - $e^{-(1+\alpha)} = (\beta h^2/2\pi m)^{3/2} n$
- so
  - $\boxed{f(p) = (\beta h^2/2\pi m)^{3/2} n e^{-\beta p^2/2m}}$
- The total energy, on the other hand, is
  - $E_{\text{tot}} = \sum_s A^s f_s \frac{p^2}{2m} \to 4\pi \left(\frac{L}{h}\right)^3 \left(\frac{\beta h^2}{2\pi m}\right)^{3/2} n \int dpp^2 \frac{p^2}{2m} e^{-\beta p^2/2m} = nL^3/\sqrt{2\pi}\beta \int dyy^4 e^{-y^2/2}$
- or, with $nL^3 = N$ and $\int dyy^4 e^{-y^2/2} = 3\sqrt{\pi/2}$,
  - $\boxed{E_{\text{tot}} = 3N/2\beta = 3Nk_{\text{B}}T/2}$
- so the mean energy per particle is $\langle E \rangle = 3/2\beta = 3k_{\text{B}}T/2$.
  - se there is $k_{\text{B}}T/2$ energy per 'translational degree of freedom'
- The entropy is
  - $S = -\sum_s A^s f_s \log f_s \to -4\pi \left(\frac{L}{h}\right)^3 (\beta h^2/2\pi m)^{3/2} n \int dpp^2 e^{-\beta p^2/2m} \log f$
- but $\log f = \log[(\beta h^2/2\pi m)^{3/2} n] - \beta p^2/2m$ here. Performing the integral of the second contribution here gives a constant – independent of temperature β^{-1} and particle density n that is – so we have
  - $\boxed{S = \text{constant} - N \log[(\beta h^2/2\pi m)^{3/2} n])}$
- This is in accord with $E_{\text{tot}} = 3Nk_{\text{B}}T/2$ – and the identification $\beta = 1/k_{\text{B}}T$ – since that implies that if we heat a fixed volume of gas we have to apply heat $dQ = dE_{\text{tot}} = 3Nk_{\text{B}}dT/2$ and the entropy is then $S_{\text{phys}} = \int dQ/T = (3Nk_{\text{B}}/2) \log T + \text{constant}$. While the statistical mechanical entropy is $S = -(3N/2) \log \beta + \text{constant}$.
- And it is physically reasonable that the entropy is solely a function of the combination $\beta^{3/2} n$. Since if we compress a volume of gas adiabatically (i.e. without adding heat) this says that the temperature will increase as $T \propto V^{-2/3}$.
  - this is in accord with what one would find by tracking the heating of the particles as they gain momentum in bouncing off the walls of the shrinking vessel
  - it is in accord with the idea that the wavelength of the de Broglie waves shrinks with the linear size of the vessel if shrunk isotropically
  - and this scaling is expressed, equivalently, as saying that the pressure P, being equal to $nk_{\text{B}}T$ varies as the -5/3 power of the volume; i.e. the 'adiabatic index' is 5/3.
- The constant in the expression for the entropy as calculated above is $\beta E_{\text{tot}} = 3N/2$
  - so it is a constant entropy *per particle*
  - the 3/2 is questionable since, at the outset, we have dropped a similar term in obtaining $S = -\sum_s A^s f_s \log f_s$, which we justified on the grounds that $|\log f_s|$ is large
  - and, on the same grounds, the constant entropy per particle is small compared to the term involving β and n
  - so we generally ignore it

A.2 Saha revisited

One can derive Saha's equation for hydrogen ionization fraction within this framework as follows:

- we consider a model where we have:
  - neutral atoms (label n) in their ground state with occupation number f_{ns}
  - positive ions (label $+$) with occupation number f_{+s}
  - electrons (label e) with occupation number f_{es}
- this is slightly over-simplified in that we are not including the various possible excited states of the neutral atoms
- the entropy is the sum over types of particles
  - $S = -\sum_s A^s(f_{ns}\log f_{ns} + f_{+s}\log f_{+s} + f_{es}\log f_{es})$
- which we will wish to maximise subject to the following constraints:
  - the total number of protons is
    * $N_n + N_+ = \sum_s A^s(f_{ns} + f_{+s})$
  - the total charge is proportional to
    * $N_+ - N_e = \sum_s A^s(f_{+s} - f_{es})$
  - and the total energy is
    * $E_{\rm tot} = \sum_s A^s(f_{ns}E_n + f_{+s}E_+ + f_{es}(E_e + \chi))$
    * where $E_n = p_n^2/2m_n = (h^2/2m_nL^2)s^2$ and similarly for E_+ and E_e and where χ is the ionisation potential
- the occupation numbers f_{ns}, f_{+s} and f_{es} that maximise S subject to their being a fixed net number of protons, zero charge, and a fixed energy are found by demanding that the variation of $S - \alpha_p(N_n + N_+) - \alpha_q(N_+ - N_e) - \beta$ with respect to, in turn, f_{ns}, f_{+s} and f_{es} vanish. This gives:

$$\begin{array}{rcl} -(1+\log f_{ns}) - \alpha_p - \beta E_n = 0 & & f_{ns} = e^{-(1+\alpha_p+\beta E_n)} \\ -(1+\log f_{+s}) - \alpha_p - \alpha_q - \beta E_+ = 0 & \Rightarrow & f_{+s} = e^{-(1+\alpha_p+\alpha_q+\beta E_+)} \\ -(1+\log f_{es}) + \alpha_q - \beta(E_e+\chi) = 0 & & f_{es} = e^{-(1-\alpha_q+\beta(E_e+\chi))} \end{array} \tag{4}$$

- and summing the occupation numbers and using $N_X = \sum_s A^s f_{Xs} \rightarrow 4\pi(L/h)^3 \int dp p^2 f_{Xs}$ for each 'species' of particles gives

$$\begin{aligned} N_n &= e^{-(1+\alpha_p)}4\pi(L/h)^3\int dp p^2 e^{-\beta p^2/2m_n} = e^{-(1+\alpha_p)}\left(\frac{\sqrt{2\pi}L}{h}\right)^3\left(\frac{m_n}{\beta}\right)^{3/2} \\ N_+ &= e^{-(1+\alpha_p+\alpha_q)}4\pi(L/h)^3\int dp p^2 e^{-\beta p^2/2m_+} = e^{-(1+\alpha_p+\alpha_q)}\left(\frac{\sqrt{2\pi}L}{h}\right)^3\left(\frac{m_+}{\beta}\right)^{3/2} \\ N_e &= e^{-(1-\alpha_q+\beta\chi)}4\pi(L/h)^3\int dp p^2 e^{-\beta p^2/2m_e} = e^{-(1-\alpha_q+\beta\chi)}\left(\frac{\sqrt{2\pi}L}{h}\right)^3\left(\frac{m_e}{\beta}\right)^{3/2} \end{aligned} \tag{5}$$

- from which we obtain

$$\frac{N_+N_e}{N_n} = \left(\frac{m_+}{m_n}\right)^{3/2}\left(\frac{\sqrt{2\pi}L}{h}\right)^3\left(\frac{m_e}{\beta}\right)^{3/2} e^{-\beta\chi} \tag{6}$$

- or, ignoring the 1-part-in-2000 difference between m_+ and m_n, and using $n_e = N_e/L^3$ gives, for the ionisation fraction, Saha's result

$$\boxed{f_{\rm ionised} \equiv \frac{N_+}{N_n} = n_e^{-1}\left(\frac{2\pi m_e}{\beta h^2}\right)^{3/2} e^{-\beta\chi}} \tag{7}$$

- where, as we found before, the pre-factor is, to order of magnitude $1/(n_e \lambda_{\rm dB}^3)$ so this gives a large enhancement if the de Broglie wavelength is small compared to the mean electron spacing $n_e^{-1/3}$.
  - since the mean squared momentum is $\langle |\mathbf{p}|^2 \rangle = 2m_e \langle E \rangle = 3m_e/\beta = 3m_e k_{\rm B} T$, the 'effective volume' in momentum space is $(\Delta p)^3 \sim (m_e/\beta)^{3/2}$ so, with $\Delta x = L$ we can write $f_{\rm ionised} \sim (N/(\Delta x \Delta p/\hbar)^3)^{-1} e^{-\beta\chi}$
  - so if the number of particles per volume $\hbar^3$ of phase space is small – the definition of non-degeneracy – the 'Saha boost factor' is large.