What a student can learn from the Saha Equation

Jayant V. Narlikar
IUCAA, Pune

1. Introduction

The seminal contribution of Meghnad Saha was his ionization equation, now well known as Saha's Ionization Equation. To say that it was an important work in astrophysics would be understating it : for, the subject of astrophysics really got going only as a result of the Saha Equation. Let us begin with an examination of this assertion. For, to a student of physics the equation provides a menu of delicious results to be enjoyed and appreciated.

Till the second decade of the last century the main observational handle on studies of stars had been their luminosities and spectra. While the luminosity could give a crude estimate of the star's distance using the inverse square law of illumination, the spectrum contained a lot more information.

For example, the continuum spectrum did, in the first approximation resemble the black body spectrum which was well known in those days. If the star was generating energy inside it and radiating it away, then it was in a state of equilibrium and provided the amount radiated was negligible compared to the total store of radiation being scattered within the star one expected the equilibrium state to resemble the black body state. This enabled the astronomer to estimate the star's surface temperature.

With surface temperatures of the order of 3000 K and above, it became clear that the matter at the surface was not likely to be in a state of neutral gas. With large thermal motions and the resulting frequent collisions, it would be impossible for the typical atoms to retain all of their orbital electrons and so they would be ionized and the matter would be in a state of plasma. How would the state of equilibrium be in such circumstances? Naturally we expect some of the atoms of the matter to remain unchanged, while some would exist as ions and free electrons. But what would be the proportions of these three ingredients? The Saha equation answered this important question by giving the following relation:

$$\frac{N_i N_e}{N} = \left(\frac{m_e \kappa T}{2\pi h^2}\right)^{3/2} \exp\left(-\frac{B}{\kappa T}\right) \quad (1)$$

Here the number N_e, N_i, N denote the number densities of free electrons, ions and neutral atoms at temperature T, B being the binding energy of the atom. The ratio of the binding energy to temperature appears in the exponential form in this equation, thus underscoring its critical effect on the equilibrium abundances of these three component species. Let us try to understand it qualitatively.

What does the binding energy indicate? Recall that an atom contains a positively charged nucleus surrounded by negatively charged electrons. The latter are held close to the nucleus by the force of electrostatic attraction. It is this force that provides the *binding* and its energy denotes what work must be done to tear the electrons apart from the nucleus. Thus the larger the binding energy B the more likely that the atom would stay intact despite any attempt at disruption.

The disruption comes from collisions of the atom by other particles. The larger the velocity $< v >$ of a colliding particle the greater the chance of a break-up of the atom. As statistical mechanics tells us, the measure of speed is through the temperature T of the syatem. The larger the temperature the greater the average velocity per particle. In fact, we know from this subject that the average kinetic energy per particle is proportional to T.

In the above equation we thus see that the larger the value of B the smaller the value of the ratio on the left hand side. That is, we will expect a smaller proportion of ions and free electrons. However, as T the temperature is raised, the right hand side increases and we get higher proportions of free ions and electrons. *In short, with rising temperature the matter moves towards the plasma state.*

In Saha's equation we therefore see the broad link between atomic physics, thermodynamics and observational astronomy. The appearance of the Boltzmann's constant κ in (1), the atomic binding energy, and the temperature indicate this tripartite relationship. This was the beginning for astrophysics : it was here that a clue was made available to interpret the spectrum of a star including the strengths of the emission and absorption lines in it in terms of the ambient state of ionization of the stellar envelope and its temperature.

The surface conditions of the star serve as valuable boundary conditions for stellar models which seek to give details of the unseen stellar interior. In the model proposed by Eddington, the star is a spherical object, made of plasma held together under its own gravitation. In fact, the inward force of gravitation can be so large that unless the star has significant pressures within, it cannot resist gravity.

Exercise: Assume a ball of pressure-free gas of the mass the same as that of the Sun, $M = 2 \times 10^{33}$g and follow its gravitational contraction from its starting radius equal to that of the Sun, i.e., $R = 7 \times 10^{10}$cm. *How long will it take to shrink to zero radius?* You will be surprised to see what small answer you get!

The classic equations of Eddington[1] are differential equations which give the march of physical quantities like density, temperature, pressure, luminous flux etc. from the centre outwards. To solve them completely the boundary conditions at the surface are required. This explains why the Saha Equation was such an important stimulus for the early astrophysics. Saha's paper[2] appeared in around 1920 and in the next 4-5 years Eddington's stellar models could be set up. That was the beginning of astronomers using the methods of laboratory physics and applying basic theories of physics to understand the large scale behaviour of stars, galaxies and the whole universe.

The purpose of this article is to emphasize the wide applicability of the Saha Equation to astrophysics : for the general impression is that the equation has relevance to stellar scenarios only. I will select two scenarios to illustrate my point, both of them far removed from stellar astrophysics. The first relates to the popular theory of the origin of the microwave background radiation in the universe and the second to the theory of the origin of light nuclei in the early universe.

2. The Microwave Background

The presently popular big bang framework of cosmology envisages the following sequence of events since the origin of the universe in a big explosion. In the early stages the universe was very hot, with typical particles of matter moving relativistically, i.e., as photons. Such a phase was said to be radiation dominated. Even electrons and protons moved with the speeds close to that of light provided the universe had a temperature of about ten thousand billion. As the universe expanded, it cooled and the speeds of particles began to drop. A crude but very useful estimate tells us that the average energy per particle is comparable to kT. Thus the speeds of the more massive particles will be lower. As the speed falls significantly below the light speed c, the particle becomes 'non-relativistic'. So in our case, first the protons become non-relativistic and then the electrons.

Standard texts in cosmology, e.g. Ref. 3, give the relevant relations describing when this happens. The later cooler epochs have the universe 'dust dominated'. That is, the universe is mainly made of matter that has negligible random motions with respect to the cosmological rest frame. Denoting the scale factor of the expanding universe by S, the simple rule is that the temperature of the radiation drops in inverse proportion to S.

As we shall see in the next section, during the period 1-200 seconds after the big bang, the temperature of the universe dropped from about 10 billion degrees to a few hundred million degrees. This was when the synthesis of nuclei took place.

The presence of nuclei, free protons, and electrons did not, however, have much effect on the dynamics of the universe, which was still radiation-dominated. But, these particles especially the lightest of them, the electrons, acted as scattering centres of the ambient radiation and kept it thermalized. The universe was therefore quite opaque to start with. For, with its frequent sctterring light could not travel in a straight line very far.

However, as the universe cooled, the Coulomb electrical attraction between the electron and the proton began to assert itself. In detailed calculations performed by P.J.E. Peebles, the mixture of electrons and protons and of hydrogen atoms was studied at varying temperatures. Because of Coulomb attraction between the electron and the proton, the hydrogen atom has a certain binding energy B. The problem of determining the relative number densities of free electrons, free protons (that is, ions), and neutral H-atoms in thermal equilibrium is therefore analogous to that we considered earlier for stars. The only difference is that the setting is cosmological rather than stellar. Following (1) we arrive at the formula relating the number densities of electrons (N_e), protons ($N_p = N_e$), and H-atoms (N_H) at a given temperature T :

$$\frac{N_e^2}{N_H} = \left(\frac{m_e \kappa T}{2\pi h^2}\right)^{3/2} \exp\left(-\frac{B}{\kappa T}\right), \qquad (2)$$

where m_e = electron mass. This equation is a particular case of *Saha's ionization equation.*

Writing N_B for the total baryon number density, we may express the fraction of ionization by the ratio

$$x = \frac{N_e}{N_B}$$

Then, since $N_H = N_B - N_e$, we get from (2)

$$\frac{x^2}{1-x} = \frac{1}{N_B}\left(\frac{m_e \kappa T}{2\pi h^2}\right)^{3/2} \exp\left(-\frac{B}{\kappa T}\right) \qquad (3)$$

For the H-atom, $B = 13.59$ eV. Substituting for various quantities on the right hand side of (3), we can solve for x as a function of T. The results show that x drops sharply from 1 to near zero in the temperature range of $\sim$ 5000 K to 2500 K, depending on the value of N_B. For example, $x = 0.003$ at $T = 3000$ K for the case where the baryon density at present is about 2×10^{-30}g cm^{-3}.

Thus by this stage most of the free electrons were removed from the cosmological brew, and as a result the main agent responsible for the scattering of radiation disappeared from the scene. The universe thus became effectively transparent to radiantion. This is called the 'epoch of last scattering'.

Thus the Saha Equation essentially fixes the epoch when radiation decoupled from matter. Subsequent to this epoch, the radiation cooled more or less undisturbed by whatever process went on with the formation of large scale structures of matter. Since it had acquired a black body spectrum prior to the epoch of last scattering, it retained that but with a steadily diminishing temperature. In fact the formula $T \propto (1/S)$ continued to hold even when radiation decoupled from matter. The microwave background we see today is that relic radiation and it should therefore carry signatures of the last scattering epoch intact. This conclusion has remarkable observational consequences since it enables us to probe the early universe by looking at the microwave background very minutely.

3. Primordial Nucleosynthesis

Let us now go further back in time to the 1-200 second epoch when the universe was hot enough for nucleosynthesis to have taken place. Here we encounter the Saha Equation in a different setting, with atomic binding replaced by nuclear binding. Free protons and neutrons could combine to form bigger and bigger nuclei provided their random speeds were slow enough for them to be trapped in the nuclear potential wells. The calculation which was first attempted by George Gamow in the late 1940s is described briefly as follows.

A typical nucleus Q is described by two quantities A = atomic mass and Z = atomic number, and is written*

$$^{A}_{Z}Q.$$

This nucleus has Z protons and $(A-Z)$ neutrons. If m_Q is the mass of the nucleus, its binding energy is given by

$$B_Q = [Zm_p + (A-Z)m_n - m_Q]c^2. \tag{4}$$

Let us now consider a unit volume of cosmological medium containing N_N nucleons, bound or free. Since the masses of protons and neutrons are nearly equal, we may denote the typical nucleon mass by m. Thus $m_n \approx mp = m$. If there are N_n free neutrons and N_P free protons in the mixture

$$X_n = \frac{N_n}{N_N}, \; X_p = \frac{N_p}{N_N} \tag{5}$$

will denote the fractions by weight of free neutrons and free protons. If a typical bound nucleus Q has atomic mass A and there are N_Q of them in our unit volume, we may denote the weight fraction of Q by

$$X_Q = \frac{N_Q A}{N_N} \tag{6}$$

Now at very high temperatures ($T \gg 10^{10}$K), the nuclei are expected to be in thermal equilibrium. However, even at these temperatures the usual

*sometimes the suffix z is suppressed.

formulae for non-relativistic thermodynamics will apply. Further, since we are now concerned with relative number densities, we need to consider the chemical potentials. Thus

$$N_Q = g_Q\left(\frac{m_Q\kappa T}{2\pi h^2}\right)^{3/2}\exp\left(\frac{\mu_Q - m_Q c^2}{\kappa T}\right) \quad (7)$$

where we have reinstated the chemical potentials μ_Q. Since chemical potentials are conserved in nuclear reactions,

$$\mu_Q = Z\mu_p + (A-Z)\mu_n \quad (8)$$

assuming that the nuclei were built out of neutrons and protons by nuclear reactions.

The unknown chemical potentials can be eliminated between (7) and similar relations for N_P and N_n, the result is expressed in this form :

$$X_Q = \frac{1}{2}g_Q A^{5/2} X_p^Z X_n^{A-Z} \times \xi^{A-1}\exp\left(\frac{B_Q}{\kappa T}\right) \quad (9)$$

where

$$\xi = \frac{1}{2}N_N\left(\frac{m\kappa T}{2\pi h^2}\right)^{-3/2} \quad (10)$$

Note that equation (9) is a reincarnation of the Saha equation (1) with nuclear binding replacing the atomic Coulomb binding!

For an appreciable buildup of complex nuclei, T must drop to a low enough value to make $\exp(B_Q/\kappa T)$ large enough to compensate for the smallness of ξ^{A-1}. This happens for nucleus Q when T has dropped down to

$$T_Q \sim \frac{B_Q}{\kappa(A-1)\mid ln\xi \mid}. \quad (11)$$

Let us consider what happens when we apply the above formula to the nucleus of ^4He. The binding energy of this nucleus is $\cong 4.3 \times 10^{-5}$ erg. If we substitute this value in (11) and estimate N_N from the presently observed value of nucleon density of around 10^{-6}cm^{-3}, we find that T_Q is as low as $\sim 3 \times 10^9$K. However, at this low temperature the number densities of participating nucleons are so low that four-body encounters leading to the formation of ^4He are extremely rare. Thus the underlying assumption of thermodynamic equilibrium (which requires frequent collisions) leading to (11) becomes invalid. We therefore need to proceed in a less ambitious fashion in order to describe the buildup of complex nuclei.

Hence, we try using two-body collisions (which are not so rare) to describe the buildup of heavier nuclei. Thus deuterium (d), tritium (^3H), and helium (^3He, ^4He) are formed via reactions like

$$\begin{aligned} p+n &\leftrightarrow d+\gamma \\ d+d &\leftrightarrow {}^3He+n \leftrightarrow {}^3H+p \\ {}^3H+d &\leftrightarrow {}^4He+n. \end{aligned} \quad (12)$$

Since formation of deuterium involves only two-body collisions, it quickly reaches its equilibrium abundance as given by

$$X_d = \frac{3}{\sqrt{2}}X_p X_n \xi \exp\left(\frac{B_d}{\kappa T}\right). \quad (13)$$

However, the binding energy B_d of deuterium is low so that unless T drops to less than 10^9 K, X_d is not high enough to start further reactions leading to ^3He, ^3H, and ^4He. In fact the reactions given in (12) with the exception of the first one do not proceed fast enough until the temperature has dropped $\sim 8 \times 10^8 K$.

Although at such temperatures nucleosynthesis does proceed rapidly enough, it cannot go beyond ^4He. This is because there are no stable nuclei with $A = 5$ or 8, and nuclei heavier than

^4He. So the process terminates there. Detailed calculations by several authros have now established this result quite firmly.

There is a fairly good agreement between these calculations and observational estimates of the light nuclei, agreement at least good enough to generate confidence in the big bang picture of a early hot universe. One could equally well argue that the success of the calculation generates confidence in the thermodynamic equilibrium picture conceived by Meghnad Saha.

4. Conclusion

I have illustrated these two examples from cosmology to emphasize the wide applicability of Saha's work today. The equation comes up when we consider the synthesis of nuclei in stars also. Indeed, Saha himself may not have imagined that this result would have important implications for cosmology. However, work of a fundamental nature in physics inevitably finds unexpected applications. The Saha Equation is an example.

References

[1] A.S. Eddington, *The Internal Constitution of the Stars* Cambridge University Press (1926)

[2] M.N. Saha, *On a Physical Theory of Stellar Spectra*, *Proc. R. Soc.*, **A99**, 135 (1921)

[3] J.V. Narlikar, *An Introduction to Cosmology* Cambridge University Press (2002)