

Physics of the 21-cm transition and the cosmological evolution of the spin temperature

Zheng Zhang*

Physics Department, Brown University

(Dated: April 20, 2019)

Abstract: We thoroughly talk about basic physics of 21 cm line. Boltzmann distribution is a good description for hyperfine states of a neutral hydrogen cloud. We derive transfer properties of 21 cm photons and Ly α photons, and the former is optically thin and the latter is optically thick. Commonly defined temperatures all have universal statistical meanings. Wouyhouysen-Field effect can also be understood as a statistical behavior intrinsically. The global history in EoR can be divided into several phases. We also discuss how to parameterize the evolution of spin temperature. One can set these parameters, f_{esc} , f_* , f_X , N_{ion} , N_α , to simulate the spin temperature.

I. INTRODUCTION

Recent years, we experienced a great leap to studying the 21 cm line. Many 21 cm observations are well processing, such as the Murchison Widefield Array(MWA), the Precision Array to Probe the Epoch of Reionization (PAPER), the 21cm Array (21CMA), Experiment to Detect the Global Epoch of Reionization Signature(EDGES) and so on. They are more and more likely to help us learn about the galaxies, the reionization and even basic physics. The goal of this paper is to go through the physics underline the 21 cm observations.

This paper is organized as follows. We first talk about the basic physics of 21 cm hydrogen transition in section II. In section III, we describe how are these signals observed. In section IV, we will talk about the global signal, including the global history of the Epoch of Reionization and a discussion on how to parameterize the evolution of the spin temperature.

II. PHYSICS OF 21 CM LINE

Top codebreakers begin their work with thoroughly surveying their opponents who invented the encryption. After then, they construct a detailed personality, and put themselves in opponents' positions. When this kind of emotional interpolation meets with the rational extrapolation, the ultimate secrets were found.

The way we decode the 21 cm signature from a neutral hydrogen cloud is nothing more than what those codebreakers do. If anything, we are lucky enough to survey our opponent in a universal way, i.e., physics. In this section, physics of atomic hydrogen and statistics of abundant hydrogen atoms will tell how the cosmic history was encrypted.

In this section, we focus on a macroscopic quantity, the ratio between the number densities n_i of the hydrogen atoms in 1S singlet and 1S triplet levels, of a neutral hydrogen cloud. (Throughout this paper, we denote 0 and 1 as singlet and triplet levels respectively.) As we will see

in following sections, it is this ratio characterized by the spin temperature T_S that determines the detectability of the 21 cm signal.

Generally, two kinds of processes cause the hyperfine transitions of hydrogen atoms and thus change the ratio n_1/n_0 , collision processes and radiation processes. We will first discuss the collision processes and introduce T_S statistically. Then we move on to the radiative processes. We will just mention the coupling of collisions, for its messy calculation but simple physics and nothing to do with observation. In contrast, we will discuss the radiative processes analytically, qualitatively and intrinsically.

A. Statistical description of an ideal neutral hydrogen cloud

Let's first discuss an ideal neutral hydrogen cloud. It's ideal for that: 1. There is no background radiation. 2. It's in an equilibrium state. We need further restricts on what does the equilibrium mean. Obviously, it doesn't only mean the kinetic equilibrium, since we are talking about statistics of the hyperfine states. The equilibrium means both kinetic equilibrium and hyperfine states (1 and 0) equilibrium. The involved degrees of freedom are kinetic degrees of freedom and hyperfine degrees of freedom. Other hyperfine states beyond 0 and 1 are reasonably not considered.

Collision and decay are two regimes in the ideal neutral hydrogen cloud regarding the 21 cm transition. However, only collision is a statistical behavior, which is the only regime of energy redistribution here in the ideal system and decay is an atomic behavior. When the efficiency of collisions is larger than that of decay, one can say this system would reach and then stay in an kinetic-and-hyperfine equilibrium state. Roughly, the spontaneous decay rate of a hydrogen atom is tiny(as discussed in the following section). In fact, to be exact, we need to solve a Boltzmann equation which couples the spin and velocity distribution[1]. But here, we just assume that a single ratio of 1 and 0 can apply to the entire hydrogen distribution.

To characterize the ratio n_1/n_0 , a macroscopic parameter, one needs to choose an appropriate statistical system. It's not a simple choice among the Fermi system,

* zheng_zhang1@brown.edu

the Bose system, or the nonexistent Boltzmann system. We must start with the intrinsic statistical mechanics—the detailed energy redistribution. In collision processes, kinetic degrees of freedom are responsible for energy exchange among not only themselves but also hyperfine freedom degrees (via $H-e^-$ collisions and $H-H$ collisions). However, the hyperfine freedom degrees could only exchange energy with kinetic freedom degrees, but not among themselves, which leads us to consider each hyperfine degree of freedom separately. For this reason, one can safely picture the environment of a hyperfine degree of freedom as: a single hyperfine freedom degree contacts with a reservoir with kinetic temperature T_k . For any given hyperfine state, a frozen single degree of freedom, the abundance of possible microstates of the reservoir gives the weight of the hyperfine state,

$$n_i \propto g_i e^{-E_i/k_B T_k}. \quad (1)$$

Thus, the ratio n_1/n_0 reads

$$\frac{n_1}{n_0} = \frac{g_1}{g_0} e^{-E_{10}/k_B T_k}. \quad (2)$$

Here one finds that the ratio follows Boltzmann distribution, but definitely not the exact well-defined Boltzmann system in textbooks. We can see that the ratio is determined by the kinetic temperature in our ideal system. Generally, we can define a parameter, the spin temperature T_S , to characterize n_1/n_0 in any system beyond equilibrium:

$$\frac{n_1}{n_0} = \frac{g_1}{g_0} e^{-E_{10}/k_B T_S}. \quad (3)$$

Finally, we come to the common definition of T_S . In our ideal case, $T_S = T_k$. We might as well always keep this model in mind where the T_S is linked to both statistical and thermal meaning.

Although we have extrapolated a statistical way to characterize the distribution of hyperfine states in an ideal neutral hydrogen cloud, we still need to learn about microscopic collisional processes. Details of atomic behaviors enable us to discuss a real system where all kinds of processes happen and one needs to figure out contributions of each source. We should mainly include (i) $H-H$ collisions; (ii) $H-e^-$ collisions; (iii) other species.

We let C_{10} and C_{01} be the de-excitation and excitation rates per atom from collisions respectively. For any type of collision, the rate must be proportional to the number of colliding particles times some function of the kinetic temperatures. Therefore, the ratio of $0 \rightarrow 1$ to $1 \rightarrow 0$ is a function of the kinetic temperature.

B. Radiative processes of 21 cm transition

If we put our ideal hydrogen cloud in a radiation field, more considerations should be taken to determine the

spin temperature. For real systems in EoR, the radiation field is CMB and ionized photons inside the IGM. There are mainly two kinds of photons involved in the hyperfine transitions between 0 and 1: the 21 cm photons and $Ly\alpha$ photons. The former accounts for direct transitions between 1 and 0, while the latter accounts for $Ly\alpha$ scattering processes, part of which significantly affects the spin temperature.

We will first describe a general radiative process, and then characterize these two corresponding processes. Throughout the following sections, we use the *specific intensity* or *brightness* I_ν to describe the energy carried by rays passing through a given direction, per unit area, frequency, solid angle and time.

1. The emission and absorption of photons by an atomic system

Before generally discussing the radiation propagation, we should first characterize the absorption and emission per atom. [citations...] Einstein first discovered the relation between emission and absorption as atomic behaviors by proposing three probable processes: spontaneous emission, stimulated emission, and absorption. They are respectively characterized by Einstein coefficients A_{21} , B_{21} , B_{12} , where 1 and 2 denote two discrete energy levels, E_1 and E_2 . A transition from 1 to 2 (2 to 1) happens by absorption (emission) of a photon of energy $h\nu_0$. Each Einstein coefficient (sec^{-1}) describes the transition probability per unit time for its corresponding process. Precisely, many processes can cause the true energy difference between E_1 and E_2 deviates from $h\nu_0$. This spectrum structure peaked at ν_0 is described by a line profile function $\phi(\nu)$. For convenience and following convention, we take the normalized line profile:

$$\int_0^\infty \phi(\nu) d\nu = 1. \quad (4)$$

Generally, $\phi(\nu)$ includes natural, thermal, pressure broadening and so on [2][3].

These Einstein coefficients are essentially atomic properties, which isn't associated with the macroscopic state of system. But we can assume the system is in detailed balance, which is a reasonable assumption since the expansion time scales promise the system in detailed balance, so that we have one more equation to limit these coefficients. Phenomenologically, the spontaneous emission is independent of brightness I_ν , but stimulated emission and absorption are proportional to I_ν . Thus, detailed balance gives the equality of numbers of transitions per unit time per unit volume out and into state 1:

$$A_{21}n_2 + B_{21}n_2I_\nu - B_{12}n_1I_\nu = 0, \quad (5)$$

which derives the brightness

$$I_\nu = \frac{A_{21}}{B_{12}(n_1/n_2) - B_{21}}. \quad (6)$$

For further discussion on relations among Einstein coefficients, we need the distribution functions for n_i , as well as the analytical form of I_ν . The detailed balance implies the hyperfine equilibrium, for which the kinetic equilibrium is a premise. As we have discussed in part A, such a atomic system follows Boltzmann distribution. Thus,

$$I_\nu = \frac{A_{21}}{B_{12}(g_1/g_2)e^{-E_{12}/k_B T} - B_{21}} \quad (7)$$

What's more, in thermodynamic equilibrium one also finds the intensity follows Plank law, which is expressed as

$$B_\nu(T) = \frac{2h\nu^3/c^2}{\exp(hc/\lambda kT) - 1}. \quad (8)$$

$I_\nu = B_\nu(T)$ gives the Einstein relations:

$$g_1 B_{12} = g_2 B_{21}, \quad A_{21} = \frac{2h\nu^3}{c^2} B_{21}. \quad (9)$$

Based on atomic properties discussed above, now we can move on to macroscopic absorption, emission and further transfer properties when a ray passes through matter. We denote the emission coefficient and the absorption coefficient as j_ν and α_ν separately. To obtain j_ν , one needs to know the frequency distribution of the emitted radiation during spontaneous decay. A good and simple assumption is that the emission is distributed in agree with the same line profile $\phi(\nu)$ that describes absorption[3].

On one hand, the emission coefficient describes the amount of energy emitted in unit volume, solid angle, frequency and time. On the other hand, each atom contributes $h\nu_0$ distributed over all directions per transition. Thus, one can express the energy of emitted radiation in two identical ways:

$$j_\nu dV d\Omega d\nu dt = (h\nu_0/4\pi)\phi(\nu)n_2 A_{21} dV d\Omega d\nu dt, \quad (10)$$

which derives

$$j_\nu = \frac{h\nu_0}{4\pi} n_2 A_{21} \phi(\nu). \quad (11)$$

Similarly, we can express α_ν , by definition $dI_\nu = -\alpha_\nu I_\nu ds$, in terms of Einstein coefficients

$$\alpha_\nu = \frac{h\nu}{4\pi} \phi(\nu)(n_1 B_{12} - n_2 B_{21}). \quad (12)$$

Note that the emission coefficient just accounts for spontaneous emission, while the stimulated emission is considered as a negative absorption into the absorption coefficient, for the merit that these two processes are both dependent on the brightness I_ν .

Thus, the radiative transfer function in terms of emission and absorption coefficients is

$$\frac{dI_\nu}{ds} = -\alpha_\nu I_\nu + j_\nu. \quad (13)$$

Optical depth is defined by

$$\begin{aligned} \tau_\nu(s) &= \int_{s_0}^s \alpha_{s'} ds' \\ &= \int ds \frac{h\nu}{4\pi} \phi(\nu)(n_1 B_{12} - n_2 B_{21}) \\ &= \int ds \frac{3c^2 A_{21}}{8\pi\nu_{21}^2} (n_1 \frac{g_2}{g_1} - n_2) \phi(\nu) \\ &= \int ds \sigma_{21} (n_1 \frac{g_2}{g_1} - n_2) \phi(\nu), \end{aligned} \quad (14)$$

here $\sigma_{21} = 3c^2 A_{21}/8\pi\nu_{21}^2$ is defined as the cross section of an atom, which is as useful as the Einstein coefficient A_{21} to characterize the possibility of decay in a gas.

We see that optical depth is measured by integrating absorption coefficient along the path of a travelling ray. $\tau_\nu > 1$ corresponds to a so called *optically thick* or *opaque* medium. When $\tau_\nu < 1$, the medium is no surprisingly said to be *optically thin* or *transparent*. The physics meaning behind these denominations seems more obvious if we express the radiative transfer equation with τ_ν and solve it. The equation can be rewritten as,

$$\frac{dI_\nu}{d\tau_\nu} = -I_\nu + S_\nu. \quad (15)$$

Here, $S_\nu \equiv j_\nu/\alpha_\nu$ is called the source function. In our considerations, the source function is constant. Then the solution reads

$$I_\nu(\tau_\nu) = I_\nu(0)e^{-\tau_\nu} + S_\nu(1 - e^{-\tau_\nu}). \quad (16)$$

2. Neutral hydrogen cloud baths in the CMB: Radiative Transfer of the 21-cm line

In this case, the wave length $\lambda = 21.1061$ cm corresponds to a frequency of $\nu = 1420.4057$ MHz. To characterize this radiative transfer we have to calculate the Einstein coefficients. One can calculate A_{10} and then use the Einstein relations to get all these coefficients so that the radiative transfer of 21 cm line is determined. We assume that $I_\gamma = I_{CMB}$, since 21 cm photons almost all come from CMB.

J. P. Wild first calculated A_{10} in his paper in 1951[4]. The transition probability of the 21-cm line is $A_{10} = 2.85 \times 10^{-15} \text{ sec}^{-1}$. Therefore one can calculate the optical depth of 21 cm line in terms of A_{10} :

$$\begin{aligned} \tau_\nu &= \int ds \sigma_{10} (1 - e^{-E_{10}/k_B T_S}) \phi(\nu) n_0 \\ &\approx \sigma_{10} \left(\frac{h\nu}{k_B T_S} \right) \left(\frac{N_{HI}}{4} \right) \phi(\nu), \end{aligned} \quad (17)$$

here N_{HI} is the column density of atomic hydrogen and the factor $\frac{1}{4}$ accounts for the ratio of atoms in the hyperfine 0 state, since all applications have $T_S \gg T^* = E_{10}/k_B$.

We can see $\tau_\nu \ll 1$, which implies the transparency of 21 cm line for transfer in neutral hydrogen cloud. This is definitely a nontrivial merit! (refer to Furlanetto 2006) For more exact expression, we use a simple assumption, which is often a good one in astrophysics, that IGM gas expands uniformly with the Hubble flow. Then the velocity broadening of a line segment s will be $\Delta V \sim sH(z)$ and the line profile function will be $\phi(\nu) \sim c/[sH(z)\nu][2]$. The column density along s can be written as $N_{HI} = x_{HI}n_H(z)s$, where x_{HI} is the neutral fraction of hydrogen. Then, an exact expression for the 21 cm optical depth reads

$$\begin{aligned} \tau_{\nu_0} &= \frac{3}{32\pi} \frac{hc^3 A_{10}}{k_B T_S \nu_0^2} \frac{x_{HI} n_H}{(1+z)(dv_{\parallel}/dr_{\parallel})} \\ &\approx 0.0092(1+\delta)(1+z)^{3/2} \frac{x_{HI}}{T_S} \left[\frac{H(z)/(1+z)}{dv_{\parallel}/dr_{\parallel}} \right]. \end{aligned} \quad (18)$$

3. Lyman lines and Wouthuysen-Field effect

In a neutral hydrogen cloud, not only the 21-cm photons play an important role in the redistribution of hyperfine states, but also Lyman photons. An atom at ground state can absorb a Lyman photon and transit to higher levels. After a while, the atom will leave the unstable state and give off a photon. Those atoms may directly decay ($nP \rightarrow 1S$) so that the atom goes back to the ground state and a Lyman photon is produced, which may change the hyperfine states. This process is then reasonably named as Lyman resonance, UV scattering or other combinations of these words. Lyman excited atoms can also cascade through intermediate levels so that different photons emerge.

Simple comparison between Ly α and higher Lyman-n levels tells us that only Ly α lines are of great influence to the redistribution of hyperfine states: The possibilities of direct decay from higher Lyman-n levels to ground state are $P_{nP \rightarrow 1S} \sim 0.8$, so typically a Lyman-n photon can scatter $1/(1 - P_{nP \rightarrow 1S}) \sim 5$ times before a decay cascade [5]. Ly α photons, by contrast, can typically scatter for hundreds of thousands times. Thus, compared with Ly α scattering, higher Lyman-n coupling is suppressed.

Siegfried Wouthuysen and George Field first explored the so called Wouthuysen-Field effect, which is illustrated in Figure 1. It says atoms can change hyperfine states through series of Ly α absorption and spontaneous re-emission. Quantum selection rules allow transitions of $\Delta F = 0, \pm 1$, except $0 \rightarrow 0$. An atomic hydrogen can jump up and down between two fine levels, 1S and 2P, but the hyperfine levels they locate are not unique. Thus, part of this scattering contributes to spin-flip.

First, Let's discuss the radiative properties of Ly α lines. As what we have done in previous sections, let's first study the radiative transfer of Ly α line to qualitatively grasp Ly α transfer. Atomic physics tell us that the

Ly α decay rate is $A_\alpha = 8\pi^2 e^2 f_\alpha / 3m_e c \lambda_\alpha^2 = 6.25 \times 10^8 \text{ s}^{-1}$, where $f_\alpha = 0.4162$ is the oscillator strength, and the frequency of Ly α line is $\nu_\alpha = 2.47 \times 10^{15} \text{ Hz}$.

One assumption is certainly permitted that the stimulated emissions can be neglected. Therefore we can express the optical depth of Lyman photons

$$\begin{aligned} \tau_\alpha &= \int dr \sigma_\alpha \phi_\alpha(\nu) n_{HI} \\ &= \frac{3A_\alpha \lambda_\alpha^3}{8\pi} \frac{x_{HI} n_H(z)}{H(z)} \\ &\approx 1.6 \times 10^5 x_{HI} (1+\delta) \left(\frac{1+z}{4} \right)^{3/2}, \end{aligned} \quad (19)$$

where we have applied assumptions on the line profile of Ly α and some tricks. More details refer to *The First Galaxies In The Universe* by Leob and Furlanetto [5]. Qualitatively, one can conclude that the neutral hydrogen cloud is optically thick for Ly α lines ($\tau_\alpha \gg 1$). These properties help us picture the scattering processes: Ly α photons travel in the neutral hydrogen cloud, not a long journey, it will be absorbed and then quickly another Ly α photon emerge. After another short journey, the same process will happen again. It can scatter for a large number of times and redistribute energy frequently, which led Wouthuysen to propose this explanation in 1952 [6]: “one can take the gas in a large container, with perfectly reflecting walls. Let the gas in equilibrium at temperature T , together with Planck radiation of the same temperature. The scattering processes will not affect the radiation spectrum.” He further reviewed that, “after a finite but large number of scattering process, the photons will obtain a statistical distribution in the vicinity of the initial frequency proportional to the Planck-radiation spectrum of temperature T ”. This is innovative but also not accurate. One can generally agree with this since it's a simple fact that, no matter whether the gas is in the container, the system is a blackbody in thermal equilibrium. But the description is not absolutely right. We will discuss the details in section D to explain that in thermal equilibrium, $T_S = T_\alpha = T_K$. Besides $T_\alpha = T_K$, we also need to care about the coupling coefficient x_α , which will be introduced and discussed in section C.

C. The spin temperature

As we have discussed, T_S is determined by several competing processes, collisional processes and radiative processes. Let's use temperatures characterize all these modes of energy. Before that we must understand the temperature exactly otherwise one will always think these temperatures are just definitions rather than intrinsic. A temperature T can be converted to $\beta = 1/k_B T$, which is a parameter of a system:

$$\beta(E) = \frac{\partial \ln \Omega}{\partial E}, \quad (20)$$

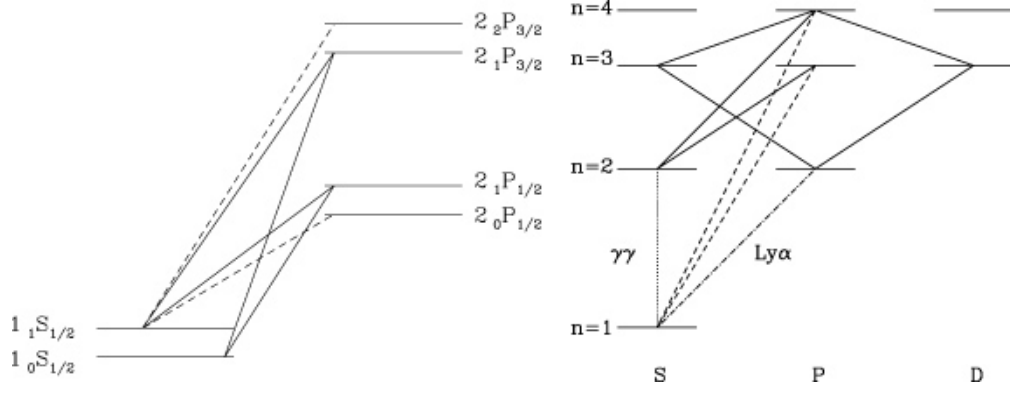


FIG. 1. Left: Level diagram illustrating the Wouthuysen-Field effect. The solid lines label transitions causing spin-flip and the dashed lines label transition uncorrelated to spin-flip. Right: Decay chains for Lyman- β and Lyman- γ excitations.[2]

where Ω is the number of microstates and E is the energy of the system. When any two systems are in thermal equilibrium, the β of two systems are equal ($\beta_1 = \beta_2$) by maximizing the number of microstates of combined system. In this paper, different modes of energy are viewed as different systems. Three temperatures are of special interest: the kinetic temperature of electrons and atoms, T_K ; the temperature of cosmic background radiation, T_γ ; and the temperature of certain lights (in the vicinity of Ly α photons), T_α .

Generally, one can characterize T_S by the equation of thermal equilibrium

$$n_1(A_{10} + B_{10}I_\nu + P_{10} + C_{10}) = n_0(B_{01}I_\nu + P_{01} + C_{01}), \quad (21)$$

where $I_\nu \approx I_{CMB}$ is the specific intensity of CMB photons.

If we assume each regime is in thermal equilibrium, we can express the ratios of transition rates in terms of their corresponding temperature of the regime. Since $T_* \ll T_K$ in collisional regime, Eqn.(2) gives

$$\frac{C_{01}}{C_{10}} = \frac{n_1}{n_0} \approx 3 \left(1 - \frac{T_*}{T_K} \right). \quad (22)$$

Similarly, B_{01}/B_{10} reads

$$\frac{B_{01}}{B_{10}} \approx 3 \left(1 - \frac{T_*}{T_\gamma} \right), \quad (23)$$

and the ratio of P_{01} and P_{10} is

$$\frac{P_{01}}{P_{10}} = \frac{n_1}{n_0} \approx 3 \left(1 - \frac{T_*}{T_\alpha} \right). \quad (24)$$

Thus, when the populations of 0 and 1 states do not change, we have[7]

$$\begin{aligned} \frac{n_1}{n_0} &= \frac{g_1}{g_0} \exp(-T_*/T_S) \approx 3 \left(1 - \frac{T_*}{T_S} \right) \\ &= 3 \frac{\frac{T_\gamma}{T_*} A_{10} + \left(1 - \frac{T_*}{T_K} \right) C_{10} + \left(1 - \frac{T_*}{T_\alpha} \right) P_{10}}{A_{10} \left(1 + \frac{T_\gamma}{T_*} \right) + C_{10} + P_{10}}, \end{aligned} \quad (25)$$

which gives

$$T_S^{-1} = \frac{T_\gamma^{-1} + x_\alpha T_\alpha^{-1} + x_c T_K^{-1}}{1 + x_\alpha + x_c}, \quad (26)$$

where

$$x_c = \frac{T_*}{T_\gamma} \frac{C_{10}}{A_{10}}, \text{ and } x_\alpha = \frac{T_*}{T_\gamma} \frac{P_{10}}{A_{10}}. \quad (27)$$

Here let's discuss more about the coupling of Lyman- α . Similar to the general radiative process we discussed above, the Wouthuysen-Field coupling must depend on the atomic properties, i.e., the total scattering rate per atom of Ly α photons,

$$P_\alpha = 4\pi\sigma_\alpha \int d\nu J_\nu(\nu) \phi_\alpha(\nu). \quad (28)$$

where $J_\nu(\nu)$ is the angle-averaged specific intensity of the radiation field. We can next relate the scattering rate P_{01} to the total scattering rate P_α . A simple assumption is that the radiation field is constant so that one can derive $P_{01} = 4P_\alpha/27$ [8]. Thus, the coupling can be expressed as

$$x_\alpha = \frac{16\pi^2 T_* e^2 f_\alpha}{27 A_{10} T_\gamma m_e c} S_\alpha J_\alpha, \quad (29)$$

where S_α is a correction of order unity, J_α is the specific flux evaluated at Ly α frequency.

D. A brief review on blackbody radiation

Blackbody is an ideal object that absorbs all incident radiation, and can also emit radiation. Radiation emitted by a blackbody in thermal equilibrium is called blackbody radiation, whose distribution follows the Planck function:

$$B_\nu(T) = \frac{2h\nu^3/c^2}{\exp(h\nu/kT) - 1}. \quad (30)$$

In fact, when a radiation is itself in thermal equilibrium, one finds the Planck law is also followed. Here in this paper, we call all those radiations following the Planck law blackbody radiation. In order to derive the Planck function as a radiation self-equilibrium distribution, let's first set up such a system. Of course, one shouldn't expect the energy distribution happens via the negligible self-interaction between photons. We need a system where photons at any frequency can freely and efficiently emerge or annihilate so that energy can be redistributed in the whole frequency domain (to be exact, one should say "permitted frequencies given by quantum mechanics", which is discrete but continuity is an excellent approximation). Since photons are massless and have 0 chemical potential, there is no number conservation law and mass is conserved all the time so that our expectation is not unreasonable. One can imagine a container with a hole. For convenience let's make it a regular box (so that we have simple periodic boundary condition). Total energy of radiation is conserved within this box and photons can be created or destroyed by the walls. One can use a grand canonical ensemble to study this system, and photons follow Bose-Einstein statistics. Thus the average number of photons of frequency ν is

$$n_\nu = \frac{1}{\exp(h\nu/kT) - 1}, \quad (31)$$

and the average energy over all states accounting for the contribution of photons with frequency ν to each state is (This long name is so stupid...I just can't come up with any other statement as accurate but more beautiful.)

$$\bar{E}(\nu) = \frac{h\nu}{\exp(h\nu/kT) - 1}. \quad (32)$$

As claimed before, we use a function with a dimension of radiation energy density (energy per volume per frequency per solid angle), u_ν , to characterize the distribution of photons. In fact, u_ν and \bar{E} describe the same thing in two different parameter spaces, the former is in "real space \otimes frequency space" and the latter is in phase space, which can be viewed as "real space \otimes k -space". Let's first figure out the equivalent integration in respective space:

$$\int d^3k = \int k^2 dk d\Omega = \int \left(\frac{2\pi}{c}\right)^3 \nu^2 d\nu d\Omega. \quad (33)$$

The shape of the container gives the boundary condition of the wave function, which derives the number of states in an element of k -space is

$$\Delta N = 2 \cdot \frac{V}{(2\pi)^3} d^3k, \quad (34)$$

here we have considered the degeneracy of photons. Thus the total number of states is

$$\begin{aligned} N &= V \int \frac{2}{(2\pi)^3} d^3k = V \int \frac{2}{(2\pi)^3} \left(\frac{2\pi}{c}\right)^3 \nu^2 d\nu d\Omega \\ &= V \int \frac{2\nu^2}{c^3} d\nu d\Omega. \end{aligned} \quad (35)$$

One finds that the density of states (the number of states per volume per frequency per solid angle) is given,

$$\rho = \frac{2\nu^2}{c^3}. \quad (36)$$

Then the diversity of expressions for the energy density derives

$$\begin{aligned} u_\nu dV d\nu d\Omega &= \bar{E} \cdot \frac{2}{(2\pi)^3} dV d^3k \\ &= \bar{E} \cdot \left(\frac{2\nu^2}{c^3}\right) dV d\nu d\Omega. \end{aligned} \quad (37)$$

Thus, we have

$$u_\nu = \frac{2h\nu^3/c^3}{\exp(h\nu/kT) - 1}. \quad (38)$$

As to the radiation intensity I_ν , which measures the flux of radiation, is achieved via $I_\nu = u_\nu c$. Thus

$$I_\nu = \frac{2h\nu^3/c^2}{\exp(h\nu/kT) - 1}, \quad (39)$$

which is exactly the Planck function.

Generally, we see that the Planck law is the one describes not only the thermal radiation of a blackbody but also the photon distributions of a radiation in thermal equilibrium. One can say they incorporate the same physics (which is discussed as the third point below), but historically, the former is a phenomenological one, while the latter is intrinsic. As for what is the "wall" in real situation, the efficient Compton scattering in the early universe is a good example. At the moment of the last scattering, the background photons were all left in good accordance to the Planck function till now, which implies the second property discussed below.

Some further discussion about the properties of blackbody radiation is helpful:

1. The temperature of a blackbody radiation can be determined by identifying either radiation intensity or the shape of spectrum (line profile) at the vicinity of a certain frequency.

2. Regimes of energy redistribution for photons: Besides the ideal wall, Compton scattering and transitions between energy levels of atom systems are all such regimes. There is no doubt that existence of these regimes is not enough for photons to reach thermal equilibrium, where the efficiency of energy redistribution should make sure the macro-properties of system are convergent in time series.

3. Considering the symmetry of absorption and emission ($B_{21} = B_{12}$) at atom scales and symmetries of other processes, one should notice that a blackbody can emit photons at any frequency when we say a blackbody can absorb all photons. From this, one can understand the blackbody as a system inside which abundant processes happen and photons at any frequency can freely emerge

or annihilate. And the thermal radiation is photons after “the last scattering”, before which countless processes have made sure a distribution with a dominant number of microstates. Therefore, it’s reasonable that the thermal radiation of a blackbody shares the same function with radiation itself in thermal equilibrium.

4. Blackbody radiation freely transferring in a space expanding with Hubble flow is always a blackbody radiation with temperature $T_1(a_1/a_2)$, where a_1 and T_1 are respectively the scale factor and temperature when the energy redistribution regime is removed and a_2 is the scale factor when the temperature is measured. A simple derivation as follows:

Denote the frequency of the same photon as ν_1 and ν_2 at a_1 and a_2 respectively. Then we have

$$a_1\nu_1 = a_2\nu_2. \quad (40)$$

Note that the just redshift happened to per photon and there is no statistical behaviors when $a_1 < a < a_2$. So the occupation number n_{ν_2} is always equal to n_{ν_1} . However, the density of states, which is of particular note, is $\rho(\nu_1) \cdot (d\nu_1/d\nu_2) \cdot (dV_1/dV_2) = \rho(\nu_1)(a_2/a_1)^{-2}$. Thus, the average energy is

$$\bar{E}(\nu_2) = \frac{h\nu_2}{\exp(h\nu_1/kT_1) - 1}, \quad (41)$$

and the energy density u_ν comes to be

$$u_{\nu_2} dV d\nu_2 d\Omega = \bar{E}(\nu_2) \left(\frac{2\nu_1^2}{c^3} \right) \left(\frac{a_2}{a_1} \right)^{-2} dV d\nu_2 d\Omega. \quad (42)$$

Thus,

$$\begin{aligned} u_{\nu_2} &= \bar{E}(\nu_2) \left(\frac{2\nu_1^2}{c^3} \right) \left(\frac{a_2}{a_1} \right)^{-2} \\ &= \frac{2h\nu_2^3/c^3}{\exp[h\nu_2/kT_1(a_1/a_2)] - 1}, \end{aligned} \quad (43)$$

which derives the brightness I_ν follows the Planck function with temperature $T = T_1(a_1/a_2)$.

5. “ $T_\alpha = T_K$ ”.

The true Ly α frequencies may be the successive vicinity of the initial Ly α frequency, or just discrete frequencies that continuity is not a good assumption. But that doesn’t matter. We know the neutral hydrogen cloud is a place where Ly α radiation can redistribute over possible frequencies. Once the intensity of a specific frequency is deviated from equilibrium value, it will receive a negative feedback(stimulated emission/absorption) so that it will be pull back to be closer to the equilibrium value. This merit of the regime together with the large number of scattering processes allow us to say the Ly α radiation can reach a thermal equilibrium so that we can use a color temperature T_α to characterize the radiation distribution. What’s more, it’s in equilibrium with the hyperfine states: in the system combined by Ly α radiation and hyperfine states, the total energy is constant.

They can exchange energy and finally the system will deposit constant energy in hyperfine system. At that time, $\beta_\alpha = \beta_S$. Besides, collisional regime allows the hyperfine states exchange energy with kinetic system, so that in thermal equilibrium $\beta_S = \beta_K$. Thus, when these regimes join together and reach a thermal equilibrium, we have $\beta_\alpha = \beta_S = \beta_K$, i.e., $T_\alpha = T_S = T_K$. But note that the Planck function incorporates the state density $\rho = 2\nu^2/c^3$ of a successive frequency domain (Although periodic boundary condition gives discrete frequencies, continuity is a good assumption.). However, since we are not sure the structure of accessible frequencies, we can’t say the shape is according to the Planck shape. Even if the shape of Ly α distribution follows the Planck shape, as pointed out in our first point, “Both magnitude and shape of the spectrum can identify the temperature of radiation”. Thus their magnitudes should be the same, not only be “proportional to Planck radiation” as Wouthuyzen said.

III. OBSERVATION

By convention, we quantify I_ν by the equivalent brightness temperature, $T_b(\nu)$, such that $I_\nu = B_\nu(T_b)$. The solution of radiative transfer equation, in the Rayleigh-Jeans limit, now can be written in a form with respect to brightness temperature,

$$T'_b(\nu) = T_{ex}(1 - e^{-\tau_\nu}) + T'_R(\nu)e^{-\tau_\nu}, \quad (44)$$

Here, the excitation temperature T_{ex} is just the spin temperature T_S , and T'_R is T_{CMB} . We hope to measure

$$\begin{aligned} \delta T_b &= \frac{T_S - T_\gamma}{1 + z} (1 - e^{-\tau_\nu}) \\ &\approx \frac{T_S - T_\gamma}{1 + z} \tau_\nu \\ &\approx 27x_{HI}(1 + \delta) \left(\frac{\Omega_b h^2}{0.023} \right) \left(\frac{0.15}{\Omega_m h^2} \frac{1 + z}{10} \right)^{1/2} \left(\frac{T_S - T_\gamma}{T_S} \right), \end{aligned} \quad (45)$$

where x_{HI} is the neutral fraction of hydrogen, δ is the fractional overdensity in baryons[9].

Then the mean/global signal is

$$\Delta T \approx 9x_{HI}(1 + z)^{1/2} \left[\frac{T_S - T_\gamma}{T_S} \right]. \quad (46)$$

IV. GLOBAL 21 CM SIGNAL

A. The global history

Next we will look at the evolution of global signal qualitatively by discussing the most important phases. This part refer to two reviews[2, 9]. Just an outline.

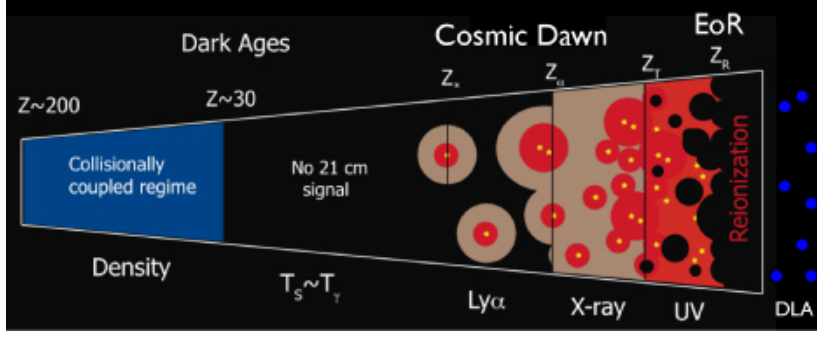


FIG. 2. Different epochs of the 21 cm signal.[9]

As shown in figure.2, some important points in time are listed. They are at redshift: $z \approx 200$, when gas and photons are about to decouple; $z \approx 30$, the beginning of the dark age; z_* , the first galaxies form; z_α , when the gas is everywhere strongly coupled to T_K ; z_h , when heating is significant and T_K is equal to T_γ again; z_T , the time of 21 cm signal saturates; and z_R , about that time Reionization completes. These epochs are not exactly determined, and even the sequence of some events are not sure.

When $200 \lesssim z \lesssim 1100$, the Compton scattering dominates the thermal equilibrium, setting $T_S = T_K = T_\gamma$. Thus there is no detectable 21 cm signal. After then, photons and the gas decouple but the gas is still very dense and collision coupling dominates, so that we still have $T_S = T_K$. Photons cool with the Hubble flow with a relationship of $T_\gamma \propto (1+z)$, while the gas follows $T_K \propto (1+z)^2$, so that there is a time, about $40 \lesssim z \lesssim 200$, T_K is less than T_γ . Thus this regime leads to an absorption global signal.

When $z_* \lesssim z \lesssim 40$, as the gas density decreases, the radiative coupling gradually plays an more important role than the collisional coupling, which sets $T_S = T_\gamma$. Thus, there is no detectable 21 cm global signal in this epoch. z_* is an important critical point when first sources emerge, which emit both Ly α photons and x-rays. In areas around these sources, Wouthuysen-Field effect set $T_S = T_K$. In this regime, Lyman- α coupling may not dominate the spin temperature as rapidly as other regimes. As more and more stars form, the Ly α coupling will saturate so that the gas will everywhere strongly couples to T_K . Thus, this epoch leaves an absorption signal, which corresponds to a $\bar{T}_K \lesssim \bar{T}_S \lesssim T_\gamma$. At $z = z_\alpha$, the critical point where Ly α flux just saturates, $\bar{T}_S = \bar{T}_K$ globally. Note that heating regime is more and more important since that time. Before the \bar{T}_K increases to be equal to T_γ , which happens at $z = z_h$, we still have the absorption signal. But after then, $T_S \sim T_K T_\gamma$, we can see emission signal. With the heating go further, the emission signal increases till $z = z_T$.

When $z_T \lesssim z \lesssim z_R$, we have $T_S \sim T_K \gg T_\gamma$. So the term of T_S in observed brightness temperature(Eqn.(45)) vanishes, i.e., $(T_S - T_\gamma)/T_S \rightarrow 1$. The magnitude of ob-

servation doesn't depend on T_S any more so that the 21 cm signal saturates. But at the same time, the filling fraction of H II regions becomes more and more significant, and finally there are just neutral hydrogen islands source the 21 cm signal[9]. All the above events are encoded in Fig.3.

B. How to parameterize the evolution of spin temperature in EoR?

Although it's not practical for the author to simulate the global history in this project. We will discuss a simple model without dramatic difference with theoretical expectation, as a practice to know more about how to characterize the reionizing universe.

From Eqn.(46), one can see that the global signal hinges on the neutral fraction of hydrogen x_{HI} and the spin temperature T_S . In the Epoch of Reionization, the spin temperature tightly couples to the kinetic temperature of gas, thus we will focus on how to express T_K .

The evolution equation of the gas temperature is

$$\frac{dT_K}{dt} = \frac{2T_K}{3n} \frac{dn}{dt} + \frac{2}{3k_B} \sum_j \frac{\epsilon_j}{n}, \quad (47)$$

The first term expresses the adiabatic cooling while the second term accounts for other heating or cooling sources, such as Compton heating, X-ray heating, Lyman- α heating and shock heating[10]. ϵ_j ($\text{erg} \cdot \text{s}^{-1} \cdot \text{cm}^{-3}$) is the density of input energy into the gas by source j . The most significant source is thought to be X-rays from quasars and first galaxies, since they can travel relatively long distance. We denote $f_{X,h}$, $f_{X,ion}$ and $f_{X,coll}$ to be the fraction of X-rays energy going to heating, ionization, and excitation respectively. For simplicity, Furlanetto assumed that the star formation rate is proportional to df_{coll}/dt , here f_{coll} is the fraction of gas collapsed onto virialized halos[2]. By this assumption, we can write

$$\frac{2}{3} \frac{\epsilon_X}{k_B n H(z)} = 10^3 K f_X \left(\frac{f_*}{0.2} \frac{f_{X,h}}{0.2} \frac{df_{coll}/dz}{0.01} \frac{1+z}{10} \right), \quad (48)$$

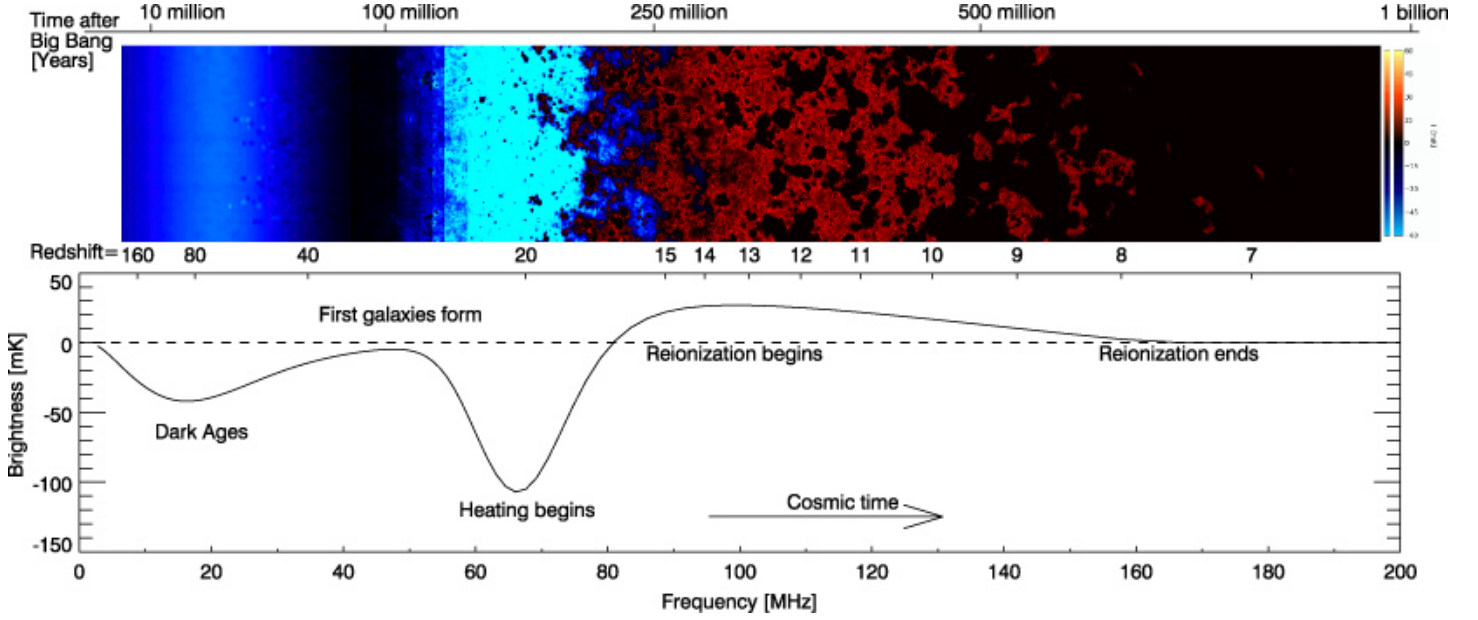


FIG. 3. The 21 cm cosmic hydrogen signal. [9] Copyright 2010 Nature Publishing Group.

where f_* is the star formation efficiency and f_X is a correction factor accounting for the difference at high redshifts. This equation tells us that X-ray heating is very rapid[2].

Next we have to solve the ionization fraction \bar{x}_i , i.e., to characterize the ionization history. Here we will associate the star formation rate (SFR) with the reionization and include both ionizing sources and recombinations[2], so

$$\frac{d\bar{x}_i}{dt} = \zeta(z) \frac{f_{coll}}{dt} - \alpha C(z, x) \bar{x}_i(z) \bar{n}_e(z), \quad (49)$$

where $\zeta = A_{He} f_* f_{esc} N_{ion}$ is the fraction of ionizing efficiency, f_{esc} is the fraction of ionizing photons that escape their host galaxy into the IGM, N_{ion} is the number of ionizing photons per baryon produced in stars, and A_{He} is

a correction for helium[10]. α is the recombination coefficient and C is the clumping factor [2][5]. One N-body simulation in the IGM gives an approximation of C [11]:

$$C(z) = 27.466 \exp(-0.114z + 0.001328z^2). \quad (50)$$

In the source term of Eqn.(49), the collapse fraction f_{coll} at high redshift is estimated as the mass fraction in halos above the cooling threshold, i.e., the minimum mass of halos where gas cools efficiently[12]. Thus, f_{coll} reads

$$f_{coll} = \text{erfc} \left[\frac{\delta_c(z)}{\sqrt{2}\sigma(m_{min})} \right]. \quad (51)$$

In order to calculate the mass, m_{min} , a virial temperature is needed, which is defined as

$$T_{vir} = \frac{\mu m_p V_c^2}{2k_B} = 1.98 \times 10^4 \left(\frac{\mu}{0.6} \right) \left(\frac{M}{10^8 h^{-1} M_\odot} \right)^{2/3} \left[\frac{\Omega_m}{\Omega_m^z} \frac{\Delta_c}{18\pi^2} \right]^{1/3} \left(\frac{1+z}{10} \right) \text{ K}, \quad (52)$$

where μ and m_p are respectively the mean molecular weight and the proton mass[12]. Roughly, the minimum mass is corresponding to a halo of $T_{vir} = 10^4 \text{ K}$. Eqn.(52) can tell us the mass corresponding to $T_{vir} = 10^4 \text{ K}$ [12]. In this way one can derive the reionization history when C and N_{ion} are set.

Further details in virial temperature need to be specified:

$$\Omega_m^z = \frac{\Omega_m(1+z)^3}{\Omega_m(1+z)^3 + \Omega_\Lambda + \Omega_k(1+z)^2}, \quad (53)$$

where Ω_k is

$$\Omega_k = -\frac{k}{H_0^2} = 1 - (\Omega_m + \Omega_\Lambda + \Omega_r). \quad (54)$$

And the final overdensity relative to the critical density at the collapse redshift is [12]

$$\Delta_c = 18\pi^2 + 82(\Omega_m^z - 1) - 39(\Omega_m^z - 1)^2, \quad (55)$$

where we have applied $\Omega_m + \Omega_\Lambda = 1$.

Now, we have expressed x_i and T_K . However, as we discussed in the global history: when $z_\star \lesssim z \lesssim z_\alpha$, there is no single source dominating the coupling. The spin temperature is between T_γ and T_K , thus we need to express the coupling coefficient x_α too, i.e., to characterize the average Ly α background (see Eqn.(29)). For estimation purpose, this quantity is approximated as[10]

$$J_\alpha \approx \frac{c}{4\pi} \bar{f}_{rec} f_\star \bar{n}_b^0 \Delta f_{coll} \frac{N_\alpha}{\delta\nu} (1+z)^2, \quad (56)$$

where \bar{f}_{rec} is the average probability that a photon in the interval (ν_α, ν_{LL}) is converted into a Ly α photon[10].

Generally, we can try to characterize the ionization history x_i , the kinetic temperature T_K and the average Lyman- α background j_α to parameterize the evolution of the spin temperature. Our extrapolation follows the below path:

$$\begin{aligned} \delta T_b &= \delta T_b(x_i, T_S) = \delta T_b(x_i, j_\alpha, T_K) \\ &= \delta T_b(x_i(\zeta, f_{coll}, C), j_\alpha, \epsilon_X) \\ &= \delta T_b(\zeta, m_{min}, f_\star, j_\alpha, \epsilon_X) \\ &= \delta T_b(f_{esc}, f_\star, f_X, N_{ion}, N_\alpha). \end{aligned} \quad (57)$$

Thus, one can simulate the global signal using a simple model and setting several parameters: $f_{esc}, f_\star, f_X, N_{ion}$, and N_α .

V. CONCLUSION

In this paper, we studied the physics of 21 cm transition, including how to characterize the distribution of the hyperfine states and radiative transfer of 21 cm line and Ly α line in neutral hydrogen gas. Also, by discussing the balckbody radiation, we explored intrinsic physics of Wouthuysen-Field effect and find that. What's more, the spin temperature sourced by competing processes and the global signal are discussed. We also derive that one can simulate the global signal using a simple model and setting several parameters: $f_{esc}, f_\star, f_X, N_{ion}$, and N_α .

ACKNOWLEDGMENTS

Thanks for Prof. Savvas. He led a great class in this semester. And thanks for Prof. Jonathan. He provided useful suggestions for this project.

-
- [1] C. Hirata and K. Sigurdson, submitted to mon. not. r. astron. soc, arXiv preprint [astroph/0605071](https://arxiv.org/abs/0605071) (2006).
 - [2] S. R. Furlanetto, S. P. Oh, and F. H. Briggs, Cosmology at low frequencies: The 21cm transition and the high-redshift universe, *Physics Reports* **433**, 181 (2006).
 - [3] G. B. Rybicki and A. P. Lightman, *Radiative processes in astrophysics* (John Wiley & Sons, 2008).
 - [4] J. P. Wild, The Radio-Frequency Line Spectrum of Atomic Hydrogen and its Applications in Astronomy., *Astrophys. J.* **115**, 206 (1952).
 - [5] A. Loeb and S. R. Furlanetto, *The first galaxies in the universe*, Vol. 21 (Princeton University Press, 2013).
 - [6] S. A. Wouthuysen, On the excitation mechanism of the 21-cm (radio-frequency) interstellar hydrogen emission line., **57**, 31 (1952).
 - [7] G. B. Field, Excitation of the hydrogen 21-cm line, *Proceedings of the IRE* **46**, 240 (1958).
 - [8] A. Meiksin, Detecting the epoch of first light in 21-cm radiation, in *Perspectives on Radio Astronomy: Science with Large Antenna Arrays* (2000) p. 37.
 - [9] J. R. Pritchard and A. Loeb, 21 cm cosmology in the 21st century, *Reports on Progress in Physics* **75**, 086901 (2012).
 - [10] S. R. Furlanetto, The global 21-centimeter background from high redshifts, *Monthly Notices of the Royal Astronomical Society* **371**, 867 (2006), <http://oup.prod.sis.lan/mnras/article-pdf/371/2/867/3458001/mnras0371-0867.pdf>.
 - [11] I. T. Iliev, G. Mellema, U.-L. Pen, H. Merz, P. R. Shapiro, and M. A. Alvarez, Simulating cosmic reionization at large scales-i. the geometry of reionization, *Monthly Notices of the Royal Astronomical Society* **369**, 1625 (2006).
 - [12] R. Barkana and A. Loeb, In the beginning: the first sources of light and the reionization of the universe, *Physics reports* **349**, 125 (2001).