THE ASTROPHYSICAL JOURNAL, 523:L1–L5, 1999 September 20

A NEW CALCULATION OF THE RECOMBINATION EPOCH

S. SEAGER,[1] D. D. SASSELOV,[1] AND D. SCOTT[2]
Received 1999 April 3; accepted 1999 July 23; published 1999 August 23

ABSTRACT

We have developed an improved recombination calculation of H, He I, and He II in the early universe that involves a line-by-line treatment of each atomic level. We find two major differences compared with previous calculations. First, the ionization fraction x_e is approximately 10% smaller for redshifts $\lesssim$800 because of non-equilibrium processes in the excited states of H. Second, He I recombination is much slower than previously thought, and it is delayed until just before H recombines. We describe the basic physics behind the new results and present a simple way to reproduce our calculation. This should enable a fast computation of the ionization history (and of the quantities such as the power spectrum of cosmic microwave background anisotropies that depend on it) for arbitrary cosmologies, without the need to consider the hundreds of atomic levels used in our complete model.

Subject headings: atomic processes — cosmic microwave background — cosmology: theory — early universe

1. INTRODUCTION

We have recently revisited the calculation of the recombination epoch of the early universe by making as few approximations as possible and by trying to retain full accuracy (Seager, Sasselov, & Scott 1999, hereafter Paper I). We were motivated by the potential to measure (with the *MAP* and *Planck* satellites) cosmic microwave background (CMB) anisotropies at the roughly 1% level over a wide range of angular scales and by indications (Hu et al. 1995, hereafter HSSW) that the existing solution for hydrogen and helium recombination has uncertainties at that level. Indeed, we uncovered a number of minor improvements as well as two more major effects that change the ionization history in a significant way.

A detailed understanding of the recombination process is crucial for modeling the power spectrum of CMB anisotropies. Since the seminal work of the late 1960s (Peebles 1968; Zeldovich, Kurt, & Sunyaev 1968), several refinements have been introduced (see discussion in HSSW and Paper I), but little has changed. Modern codes for evolving the ionization fraction $x_e = n_e/n_{\rm H}$ (where n_e is the number density of electrons and $n_{\rm H}$ is the total number density of H nuclei) have been based almost entirely on the single differential equation introduced 30 years ago, with a more accurate recombination coefficient but no other basic improvement.

With today's computing power, there is not the necessity for making sweeping approximations that existed 30 years ago. We believe our work represents the most accurate picture to date of how exactly the universe as a whole became neutral. In this Letter, we summarize what is new in the physics involved, and we present an approximate treatment of our set of equations that accurately reproduces our complete calculation. We discuss in detail the physical basis for each approximation and pay attention to the limits of validity. This Letter is supplemented by a computer code (RECFAST) that can be used to do the calculation and, in conjunction with, e.g., CMBFAST by Zaldarriaga, Spergel, & Seljak (1997), to compute accurate CMB power spectra for different cosmologies.

[1] Department of Astronomy, Harvard University, 60 Garden Street, Cambridge, MA 02138.
[2] Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC, V6T 1Z1, Canada.

2. HOW THE UNIVERSE BECAME NEUTRAL

2.1. *Our Multilevel Calculation*

In the canonical hot big bang picture, the recombination epoch is when the universe became cool enough for protons to capture electrons and form neutral hydrogen. This recombination process was not instantaneous because the electrons, captured into different atomic energy levels, could not cascade instantaneously down to the ground state. The electrons were impeded because of the fast reionizations out of excited states that were due to the huge reservoir of low-energy photons and because of the high optical depth of the Lyman lines and the continuum transitions to the ground state. Any Lyman line or continuum transition to the ground state emitted a photon with energy in which there were few blackbody photons, and this immediately photoexcited or photoionized a neighboring atom in the ground state. Atoms reached the ground state either through the cosmological redshifting of the Lyα line photons or by the 2*s*–1*s* two-photon process. Because these rates from $n = 2$ to the ground state were much slower than the net recombination rate to $n = 2$, a "bottleneck" occurred that slowed down the entire recombination process. The universe expanded and cooled faster than recombination could be completed, and a small fraction of free electrons and protons remained. This fraction, during and after recombination, affects the CMB anisotropies through the precise shape of the thickness of the photon last scattering surface (i.e., the visibility function).

The "standard" methodology considers an "effective three-level atom" with a ground state, first excited state ($n = 2$), and continuum, with the $n > 2$ states represented by a recombination coefficient. A single ordinary differential equation (ODE) can then be derived to describe the ionization fraction (eq. [1]; see Peebles 1968, 1993). Many assumptions go into this derivation, including the following: that H excited states are in equilibrium with the radiation; that stimulated de-excitation is negligible for the Lyα transition; that a simple recombination coefficient can be used; that every net recombination results in a ground-state atom, so that the ground-state number density $n_1 = n_{\rm H} - n_p$ (where $n_{\rm H}$ is the total hydrogen number density, ionized or neutral, while n_p is the proton number density, i.e., ionized H); that the Lyα redshifting can be dealt with using a simple escape probability; that collisional processes are negligible; and that He can be ignored.

Our new methodology, which is made possible with modern computing power, is to calculate recombination with as few approximations as possible. Instead of evolving a single ODE for dx_e/dt for H recombination, we evolve one ODE for each of the 300 atomic energy level populations, as well as one for electrons and one for the matter temperature. With this level-by-level treatment, we include thousands of bound-bound transitions that correctly couple the ODEs, and we calculate recombination to each atomic level as it evolves with redshift, which eliminates the need for an effective recombination coefficient (α in § 3). We also included other effects (which turned out to be negligible), such as collisional transitions, feedback of distortions to the radiation field, complete heating and cooling terms, and H chemistry (see Paper I for full details on this method). Our new method tests the approximations used in deriving the standard ODE, and we find all of them to be valid, except for the equilibrium assumption. The causes and consequences of this are discussed in § 2.2.1, and the effect can be approximated using the single ODE approach, with a small modification that artificially speeds up the recombination at low z, as discussed in § 3.

With larger ionization potentials, He I and He II recombined before H. He II recombination shows no deviation from previous calculations in that we find only small differences compared with the Saha equilibrium. However, He I recombination is important for the CMB anisotropies (HSSW) and, to some extent, for the chemistry in the early universe (Stancil et al. 1999), and this requires closer scrutiny. Our code evolves He I and He II in the same level-by-level method described above and simultaneously with H. Previous calculations for He I used the Saha equilibrium method (e.g., HSSW) or an effective three-level method (Matsuda, Sato, & Takeda 1971). We find that the excited states of He I remain in equilibrium with the radiation (unlike H) but that accurate treatment of He I results in a much delayed recombination. This can also be reproduced using a simple ODE for He I.

2.2. *What is New in the Physics?*

2.2.1. *Hydrogen*

Neutral H production is slightly faster when using our multilevel calculation compared with the standard calculation (updated with the most recent parameters [HSSW]), as shown in Figure 1.

For the standard equilibrium case, the net bound-bound rates are zero, and this is an implicit assumption in deriving equation (1). We find that at $z \lesssim 1000$, the net bound-bound rates become different from zero because, at low temperatures, the cool blackbody radiation field means that there are few photons for photoexcitation of high-energy transitions (e.g., 70–10, 50–4, etc.). In this case, spontaneous de-excitation dominates, causing a faster downward cascade to the $n = 2$ state. In other words, once an electron is captured at, say, $n = 70$, it can cascade down to the $n = 2$ state faster than in the equilibrium case because few photons are around to photoexcite it. In addition, the faster downward cascade rate is faster than the photoionization rate from the upper state, and one might view this as radiative decay stealing some of the depopulation "flux" from photoionization. Both the faster downward cascade and the lower photoionization rate contribute to the faster net recombination rate.

An imbalance develops as the background radiation continues to cool while the entire Lyman series remains optically thick (the bottleneck gets worse). A Boltzmann distribution

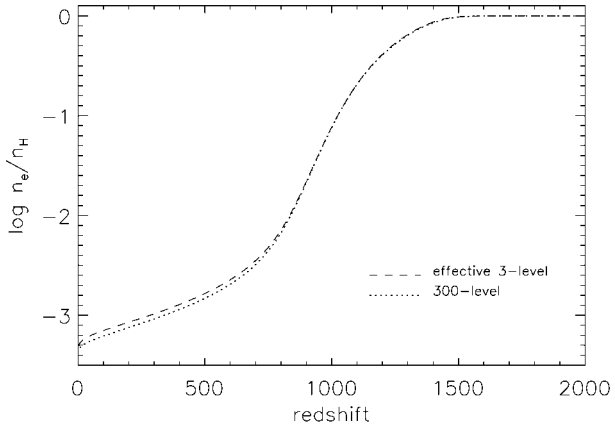

FIG. 1.—Multilevel hydrogen recombination for the standard cold dark matter (CDM) parameters $\Omega_{tot} = 1.0$, $\Omega_B = 0.05$, $H_0 = 50$, $Y_P = 0.24$, and $T_0 = 2.728$ K. The complete calculation of Paper I is represented by the dotted line, and the standard (effective three-level atom) calculation is represented by the dashed line.

relative to $n = 2$ is no longer sustainable—the excited states become progressively overpopulated. (Note that this is not a population inversion.) The radiation field is cool but strong, and it is able to keep neighboring states well coupled and to hold the highest Rydberg states in equilibrium with each other. However, in comparison with the standard equilibrium capture-cascade calculation for α, this unusual situation leads us to higher effective recombination rates for the majority of excited states without increasing photoionization proportionally. This results in a higher net rate of production of neutral hydrogen atoms. Because the net effect of our new H calculation is a faster recombination (a lower freezeout ionization fraction), our results can be reproduced by artificially speeding up recombination in the standard calculation, simply by multiplying the recombination and ionization coefficient by a "fudge factor" F.

2.2.2. *Helium*

Our new multilevel calculation shows a significantly delayed He I recombination compared with previous calculations. In fact, for our low-Ω_B models, He I recombination is still finishing at the very beginning of H recombination. Figure 2 shows the ionization fraction x_e through He II, He I, and H recombination, plotted against the standard H calculation and the He II and He I recombination calculated using the Saha ionization equilibrium equation.

The recombination of He I is slow for the same reasons that H recombination is—namely, because of the optically thick n^1p–1^1s transitions that make cascades to the ground state slow and because of the exclusion of recombinations to the ground state. In other words, He I follows a case B recombination with an inhibition factor. Because the bottleneck at $n = 2$ largely controls recombination, it is not surprising that He I and H recombination occur at a similar redshift; the ionization energy of $n = 2$ is similar in both.

The physics of He I recombination is different from that for H recombination because of its different atomic structure. More specifically, the high excited states of He I are much more strongly packed toward the continuum compared with those of H: the energy difference between the $3p$ levels and the continuum is 1.6 eV for He I versus 1.5 eV for H, compared with

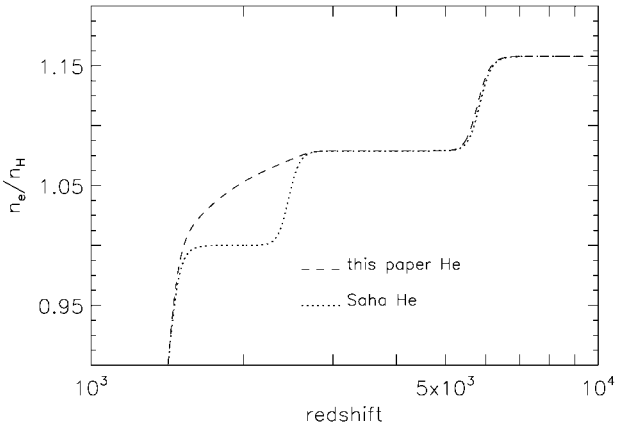

FIG. 2.—Helium and hydrogen recombination for the standard CDM parameters with $Y_p = 0.24$ and $T_0 = 2.728$ K. The first step, from right to left, is the recombination of He III to He II, and the second step is He II to He I. The dashed line represents our new He I calculation, while the dotted line represents the He Saha equilibrium recombination and the H standard calculation.

24.6 eV versus 13.6 eV for the ground state. Because of this, the radiation field has a stronger effect on the excited states of He I than those of H. This has three main consequences: (1) The strong radiation field during He I recombination keeps the ratio of the photoionization rate to the downward cascade rate higher than in the H case, resulting in a slower recombination than H. (2) The strong radiation field also causes the triplet states to be virtually unpopulated. The triplets have only forbidden (i.e., very slow) transitions with the singlets (e.g., n^3p–n^1s) and a metastable "ground state" that is 4.77 eV from the continuum. (3) The radiation field is strong enough that the excited states remain in equilibrium with the radiation throughout recombination.

One possibility for speeding up the slow He I recombination would be the existence of some neutral H that could "steal" He I resonance-line photons and invalidate the effective case B by removing the bottleneck at $n = 2$ and making it a Saha equilibrium recombination instead. However, our detailed calculation shows that the rate for this process never comes close to being significant.

Because He I recombination simply follows a case B recombination and, unlike for H, the excited singlet states remain in equilibrium throughout recombination, it can be reproduced using the effective three-level single ODE (shown below). Because the triplet excited states remain unpopulated during recombination, the effective three-level atom is composed of the ground state 1^1s, the first singlet excited state, and the continuum.

3. THE APPROXIMATE SET OF EQUATIONS

Here we present a set of equations that allow a simple way to reproduce our new recombination results from the multilevel code. Although approximate, the set provided below includes more of the recombination physics than the standard calculation, and it reproduces approximately (via parameterizations) the departures from equilibrium in H and the slow He I recombination found in Paper I.

However, a word of caution is necessary: the set of equations below has ranges of validity that correspond roughly to the choice of typical cosmologies discussed in Paper I. Attempts to calculate recombination in more extreme cosmologies, or with the addition of extra physics, will probably require the evolution of the full set of equations of the multilevel code. In addition, our approximations here are designed for optimal use with the CMB anisotropies and do not fit as well the range below $z \leq 300$, where molecular formation becomes important. Therefore, a detailed study of the chemistry in the early universe will probably also require the evolution of the full set of equations of the multilevel code.

The approximate set of equations below—two ODEs for the ionization fractions of H and He I and an equation for the matter temperature—should be solved simultaneously. We recommend leaving out He II recombination entirely since it has no effect on the power spectrum of CMB anisotropies. Otherwise, the Saha equilibrium is an adequate approximation. The two ODEs are derived from a consideration of the detailed balance in the effective three-level atoms of H and He I (see Peebles 1968, 1993). The total recombination coefficients in each case have been parameterized to reproduce our multilevel results. The equation of the total rate of change of the matter temperature is a truncated version of equation (69) in Paper I, including adiabatic and Compton cooling terms. The set is as follows:

$$\frac{dx_p}{dz} = \frac{[x_e x_p n_{\rm H} \alpha_{\rm H} - \beta_{\rm H}(1 - x_p)e^{-h\nu_{\rm H\,2s}/kT_M}][1 + K_{\rm H}\Lambda_{\rm H} n_{\rm H}(1 - x_p)]}{H(z)(1 + z)[1 + K_{\rm H}(\Lambda_{\rm H} + \beta_{\rm H})n_{\rm H}(1 - x_p)]}, \tag{1}$$

$$\begin{aligned}\frac{dx_{\rm He\,II}}{dz} = \{&[x_{\rm He\,II} x_e n_{\rm H} \alpha_{\rm He\,I} - \beta_{\rm He\,I}(f_{\rm He} - x_{\rm He\,II})e^{-h\nu_{\rm He\,I\,2^1s}/kT_M}] \\ &\times [1 + K_{\rm He\,I}\Lambda_{\rm He} n_{\rm H}(f_{\rm He} - x_{\rm He\,II})e^{-h\nu_{\rm He\,I\,2^1p2^1s}/kT_M}]\}\Big/ \\ &\{H(z)(1 + z)[1 + K_{\rm He\,I}(\Lambda_{\rm He} + \beta_{\rm He\,I})n_{\rm H} \\ &\times (f_{\rm He} - x_{\rm He\,II})e^{-h\nu_{\rm He\,I\,2^1p2^1s}/kT_M}]\},\end{aligned} \tag{2}$$

$$\alpha_{\rm H} = F10^{-19}\frac{at^b}{1 + ct^d}\ {\rm m^3\ s^{-1}}, \tag{3}$$

$$\alpha_{\rm He\,I} = q\left[\sqrt{\frac{T_M}{T_2}}\left(1 + \frac{T_M}{T_2}\right)^{1-p}\left(1 + \frac{T_M}{T_1}\right)^{1+p}\right]^{-1}\ {\rm m^3\ s^{-1}}, \tag{4}$$

$$\frac{dT_M}{dz} = \frac{8\sigma_{\rm T} a_R T_R^4}{3H(z)(1 + z)m_e c}\frac{x_e}{1 + f_{\rm He} + x_e}(T_M - T_R) + \frac{2T_M}{(1 + z)}. \tag{5}$$

Equation (5) is for the matter temperature, which we recommend using in the entire calculation above, because of the small but important effect resulting from the difference between T_M and T_R at low redshift. Below we list all constants and parameters that appear in the above equations.

The constants are Boltzmann's constant k, Planck's constant h, the speed of light c, the Thomson scattering cross section

σ_T, the electron mass m_e, and the radiation constant a_R. The three independent variables are the proton fraction $x_p = n_p/n_H$, the singly ionized helium fraction $x_{He\,II} = n_{He\,II}/n_H$, and the matter temperature T_M. The dependent variable is the electron fraction $x_e = n_e/n_H = x_p + x_{He\,II}$. Here n refers to the number density, and n_H is the total hydrogen number density.

Turning to the atomic data, the H Lyα rest wavelength is $\lambda_{H\,2p} = 121.5682$ nm. The H $2s$–$1s$ frequency, $\nu_{H\,2s} = c/\lambda_{H\,2p}$, is close enough to Ly$\alpha$ that the same averaged wavelength value can be used. The He I 2^1p–1^1s wavelength is $\lambda_{He\,I\,2^1p} = 58.4334$ nm. Note that the He I 2^1s–1^1s frequency is $\nu_{He\,I\,2s} = c/60.1404$ nm. Unlike for H, the separation of He I 2^1p and 2^1s is large enough that $\lambda_{He\,2^1p}$ and $\lambda_{He\,2^1s}$ must be distinguished, hence the extra exponential term in equation (2) over equation (1), with $\nu_{He\,I\,2^1p2^1s} = \nu_{He\,I\,2^1p} - \nu_{He\,I\,2^1s}$. The H $2s$–$1s$ two-photon rate is $\Lambda_H = 8.22458\ s^{-1}$ (Goldman 1989), while the He I 2^1s–1^1s two-photon rate is $\Lambda_{He} = 51.3\ s^{-1}$ (Drake, Victor, & Dalgarno 1969).

Furthermore, α_H is the case B recombination coefficient for H from Hummer (1994) and fitted by Péquignot, Petitjean, & Boisson (1991), with $a = 4.309$, $b = -0.6166$, $c = 0.6703$, $d = 0.5300$, and $t = T_M/10^4$ K. The factor F is 1.14 and allows equation (1) to agree with our multilevel calculation by speeding up recombination. Note that it also enters into the coefficient β via α as described below. And α_{He} is the case B He I recombination coefficient for singlets from Hummer & Storey (1998). The parameters are $q = 10^{-16.744}$, $p = 0.711$, $T_1 = 10^{5.114}$ K, and T_2 fixed arbitrarily at 3 K. This fit is good to better than 0.1% over the relevant temperature range (4000–10,000 K) and is still fairly accurate over a much wider range of temperatures. The β's are photoionization coefficients and are calculated from the recombination coefficients by $\beta = \alpha(2\pi m_e kT_M/h^2)^{3/2} \exp(-h\nu_{2s}/kT_M)$. Here ν_{2s} and α are different for H and He I. Note that T_M and ν_{2s} are used here and that incorrectly using T_R or ν_{2p} will cause a small but important difference for high baryon models.

The cosmological parameters are the redshift z, the Hubble factor $H(z)$, the cosmological redshifting of H Lyα photons $K_H \equiv \lambda_{H\,2p}^3/[8\pi H(z)]$, the cosmological redshifting of He I 2^1p–1^1s photons $K_{He\,I} \equiv \lambda_{He\,I\,2^1p}^3/[8\pi H(z)]$, and the radiation temperature $T_R = T_0(1+z)$. The primordial He abundance was taken to be $Y_p = 0.24$ by mass (Schramm & Turner 1998), and the present-day CMB temperature T_0 was taken to be 2.728 K (the central value determined by the FIRAS experiment; Fixsen et al. 1996).

A word of caution about the numerical computation. The equations are generally stiff, and so there are two approaches that can be taken. One is to use an integrator appropriate for stiff sets of equations, the other is to use Saha equilibrium values of x_e to carry the integrations through the stiffest regimes (the beginning of each recombination epoch). The former case is not much slower and certainly will work for arbitrary cosmology. For the latter case, some experimentation may be necessary in order to choose efficiently the redshift to start and finish the Saha approximation for particular cosmologies, and then it may be faster.

For He II, it is sufficient to use the relevant Saha equation:

$$\frac{(x_e - 1 - f_{He})x_e}{1 + 2f_{He} - x_e} = \frac{(2\pi m_e kT)^{3/2}}{h^3 n_H} e^{-\chi_{He\,II}/kT}. \quad (6)$$

For the nonstiff integrator approach, this can be switched on at, say, $z > 5000$. Then there will be a period when He is all singly ionized, and after, say, $z = 3500$, the Saha equation for He I needs to be switched on:

$$\frac{(x_e - 1)x_e}{1 + f_{He} - x_e} = 4\frac{(2\pi m_e kT)^{3/2}}{h^3 n_H} e^{-\chi_{He\,I}/kT}. \quad (7)$$

This can be used until, say, 1% of the He I has recombined, at which point the coupled ODEs can be solved using a routine such as DVERK.[3] It is also worth fixing the hydrogen so as to be fully ionized until some redshift and then using the Saha equation for H to get recombination started. It may also save some integration time if He recombination is switched off once the fraction of neutral He has fallen below some small number.

We found that in fact it was not necessary to evolve the H and He equations simultaneously, at least for cosmologies that we checked in detail. It appears to be sufficient to evolve each separately, even in cases where there is a small overlap in the recombination epochs. However, there is little computational expense in solving the three ODEs simultaneously, and so that is what we recommend.

4. CONCLUSION

We have presented the basic physics behind our improved recombination calculation that shows a significantly delayed He I recombination and a 10% lower residual x_e at freezeout compared with previous calculations. We present a set of equations for reproducing our work, and these equations are modified versions of those previously used because our new, detailed calculation agrees very well with the results of the standard calculation. This underscores the tremendous achievement of P. J. S. Peebles, B. Ya. Zeldovich, and colleagues in so fully understanding cosmic recombination 30 years ago. However, the goal of modern cosmology is to determine the cosmological parameters to an unprecedented level of precision, and in order to do so, it is now necessary to understand very basic things, like recombination, much more accurately.

We would like to thank George Rybicki, Ian Dell'Antonio, Avi Loeb, and Han Uitenbroek for many useful conversations. We would also like to thank David Hummer and Alex Dalgarno for discussions on the atomic physics and Jim Peebles for discussions on several aspects of this work. We thank Martin White, Wayne Hu, and Uroš Seljak for insisting on this Letter and the referee for helpful comments. D. S. is supported by the Canadian Natural Sciences and Engineering Research Council.

[3] See http://www.cs.toronto.edu/NA/dverk.f.gz by T. E. Hull, W. H. Enright, and K. R. Jackson (1976).

REFERENCES

Drake, G. W. F., Victor, G. A., & Dalgarno, A. 1969, Phys. Rev., 180, 25
Fixsen, D. J., Cheng, E. S., Gales, J. M., Mather, J. C., Shafer, R. A., & Wright, E. L. 1996, ApJ, 473, 576
Goldman, S. P. 1989, Phys. Rev. A, 40, 1185
Hu, W., Scott, D., Sugiyama, N., & White, M. 1995, Phys. Rev. D, 52, 5498 (HSSW)
Hummer, D. G. 1994, MNRAS, 268, 109
Hummer, D. G., & Storey, D. J. 1998, MNRAS, 297, 1073

Matsuda, T., Sato, H., & Takeda, H. 1971, Prog. Theor. Phys., 46, 416
Peebles, P. J. E. 1968, ApJ, 153, 1
———. 1993, Principles of Physical Cosmology (Princeton: Princeton Univ. Press), chap. 6
Péquignot, D., Petitjean, P., & Boisson, C. 1991, A&A, 251, 680
Schramm, D. N., & Turner, M. S. 1998, Rev. Mod. Phys., 70, 318
Seager, S., Sasselov, D. D., & Scott, D. 1999, ApJ, submitted (Paper I)
Stancil, R., et al. 1999, in preparation
Zaldarriaga, M., Spergel, D. N., & Seljak, U. 1997, ApJ, 488, 1
Zeldovich, Ya. B., Kurt, V. G., & Sunyaev, R. A. 1968, Zh. Eksp. Teor. Fiz., 55, 278 (English transl., Soviet Phys.—JETP, 28, 146 [1969])