PART I

SPECTRAL LINE INVESTIGATIONS

PAPER I

STUDIES OF THE 21-CM. LINE AND THEIR INTERPRETATION

INTRODUCTORY LECTURE BY

H. C. VAN DE HULST
University Observatory, Leiden, Netherlands

1. HISTORY

The subject of the 21-cm. line needs no introduction to this audience. Perhaps as an introduction I may mention one moment that belongs to the pre-history of the investigations of this line. In the spring of 1944 Oort said to me: 'We should have a colloquium on the paper by Reber; would you like to study it? And, by the way, radio astronomy can really become very important if there were at least one line in the radio spectrum. Then we can use the method of differential galactic rotation as we do in optical astronomy.'

This hope has been amply fulfilled. You know the further history: 1945, first mention of the 21-cm. line in the literature on radio astronomy; 1947, laboratory measurements of its frequency; 1951, its discovery in the Galaxy; 1952, first plots of spiral structure over a good part of the galactic plane; 1953, detection of hydrogen radiation from the Magellanic clouds; 1954, discovery of high-velocity wings in the central regions; 1954, discovery of absorption effects.

2. BASIC PROBLEMS

The basic physical data seem to be complete. The statistical weights of upper and lower level are 3 and 1. The inverse transition probability is 11 million years; the rest frequency is 1420·4056 Mc./s.; there are no significant broadening effects besides Doppler effect.

The basic astronomical data are not so well known but must be found from the observational work on the 21-cm. line and from other astronomical sources. These unknowns are:

(*a*) The density of the atomic hydrogen gas.

(*b*) The state of motion of the gas (systematic and random motions).

(*c*) The temperature of the gas.

In practice it is impossible to decide how all these vary from place to place in the Galaxy. Therefore, simplifying assumptions are made, e.g. the assumption of homogeneous temperature and random motions, in order to find the density distribution (van de Hulst, Muller and Oort[1]).

3. TEMPERATURE AND OPTICAL DEPTH

The temperature occurring in the formulae is the temperature that describes the precise population ratio of the upper and lower levels. For this temperature determines how much smaller the effective (classical) absorption coefficient is than the quantum absorption coefficient. The formulae are well-known. Let the index n denote the upper level and the index m the lower level. Then the population ratio

$$N_n/N_m = 3e^{-h\nu/kT} = 3\left(1 - \frac{h\nu}{kT}\right). \tag{1}$$

Let B_{mn} and B_{nm} be the transition probabilities for absorption and spontaneous emission. The classical absorption coefficient is proportional to

$$N_m B_{mn} - N_n B_{nm} = N_m B_{mn}\left\{1 - 3\left(1 - \frac{h\nu}{kT}\right)\frac{1}{3}\right\} = N_m B_{mn}\frac{h\nu}{kT}. \tag{2}$$

Precisely, this quantity is $4\pi/hc$ times the classical absorption coefficient per unit length, integrated over the frequencies in the broadened line. Upon multiplication with the incident intensity, averaged over all solid angles, this gives the effective number of *absorption transitions* per unit volume per second.

As a check we may multiply it by the classical intensity of Rayleigh-Jeans at the same T. By Kirchhoff's law the product must be also the number of *emission transitions* per unit volume per second:

$$N_m B_{mn}\frac{h\nu}{kT}\cdot 2\nu^2 c^{-2} kT = 3N_m \cdot \tfrac{1}{3}B_{mn}\frac{2h\nu^3}{c^2} = N_n A_{nm}. \tag{3}$$

It is important to note that the dependence on T cancels, as it should, for the radiation occurs by spontaneous transition.

The consequence is that *only* absorption measurements of the 21-cm. line can give information about this temperature. Self-absorption (saturation-effects) or absorption of continuous radiation will do, in principle.

Emission measurements on masses of gas of small optical depth, as we see in most directions, *cannot* tell us anything about the temperature.

The temperature defined above may be identified for most practical purposes with the kinetic temperature of the gas, which is defined by the Maxwell velocity distribution of the atoms. This follows from the computations of Purcell and Field[2] on the number of exchange collisions, by which radiationless transitions between the two levels occur.

So far we have made only the assumption that a local temperature exists. The determination of the temperature is simple if we make the additional assumption that the temperature is homogeneous, i.e. equal in all volume elements along the line of sight. It requires the knowledge of the saturated intensity in absolute units. The saturated intensity may be found by a rough curve-of-growth computation, if we know the intensity at two places with a very different but known ratio of the optical depths. This ratio may be computed fairly reliably from the differential galactic rotation. The absolute scale follows from a calibration of the antenna temperature, which is a separate problem. A numerical example from the Leiden data is:

highest antenna temperature in survey (at $l = 43\overset{\circ}{.}2$, $b = -1\overset{\circ}{.}0$):

$$T = 118^\circ,$$

reduced to zero band-width and beam-width:

$$T = 124^\circ,$$

estimated optical depth $\tau = 4{\cdot}2$, so gas temperature:

$$T_0 = 126^\circ.$$

The DTM observers find a higher value, $T = 150^\circ$. The preliminary Australian data gave even higher peaks in some directions. This may be due to calibration uncertainties or to local temperature differences.

A serious indication for the existence of local temperature differences is that an observed temperature of this order cannot easily be reconciled with the theoretical temperatures for H I regions of Spitzer and Savedoff[3] which are near 50°. The factor is only 2 or 3 but the heat budget is drawn out of balance by several powers of ten. Kahn[4] has indicated heating by collision of clouds as the obvious extra gain; this comes only occasionally and after it the gas cools down gradually. He also has shown that the measured saturation temperature T_0 is the harmonic mean of the actual temperatures (weighted with numbers of atoms in the

line of sight) provided there is a rapid succession of layers with different temperatures within any small optical depth. The latter assumption is astronomically not plausible. It is more plausible that in some directions the clouds nearest the sun are thick enough to impress their temperature strongly on the saturation value. But the directions and frequencies in which we might check this are not too numerous. This is another important reason for stressing the need of accurately comparable calibrations in all regions of the sky.

A modification of the reduction is needed if there is a continuous background to the spectral line. The method to be used depends on the location and properties of the sources of the continuous radiation. Let us assume that the instrument records the intensity difference of the radiation at a frequency inside the line and outside the line.* The formulae (section 7) become simple if we make the assumption that the gas density is homogeneous across the beam. The recorded intensity difference is then unaffected by sources of continuous radiation that are nearer to us than the hydrogen gas. It is $(I_0 - I_c)\,(1 - e^{-\tau})$, instead of $I_0(1 - e^{-\tau})$, if the sources are beyond the gas. Here I_0 is the saturation intensity of the gas radiation, I_c the continuous intensity of the sources as measured by this antenna and τ the optical depth of the gas. Only if the sources are beyond the gas a modified reduction is needed. As the absorption measurements have already shown that the assumption of homogeneity across the beam is not fulfilled we have to use this modified reduction method with caution. Fortunately, until now the continuous spectrum necessitated only a small correction.

4. SURVEYS

Many observational results reported in this symposium have the form of surveys. They are approximations to the ideal of a scan of the intensity in three dimensions: longitude, latitude and frequency. From it we should like to obtain the hydrogen intensity distribution as a function of four dimensions: longitude, latitude, distance and velocity. The data obviously cannot be sufficient. So a simplifying assumption is made, namely that the velocity distribution is known at each point. It is assumed to be centred on the velocity corresponding to circular motion in the Galaxy and to be spread by a known distribution of random cloud velocities. This assump-

* This is not precisely true in Muller's receiver which operates on an A.V.C. (automatic volume control) system but, as the instrumental noise level is high, the measured noise ratio exceeds the value 1 by a small amount. This amount is recorded and is very nearly proportional to the difference that is sought.

tion makes a full reduction possible, at any rate for the small latitudes, $b < 15°$. This reduction method used in the Leiden surveys will be called the *standard method*.

5. THE STANDARD METHOD AND ITS LIMITATIONS

There are blurring effects in all three observational co-ordinates, expressed by beam-width and band-width. Additional blurring is caused by the distribution of cloud velocities in converting from frequencies to distance. In the Leiden survey the distance between half-power points in the blurring function in the Perseus arm at $l = 90°$, $r = 2{\cdot}8$ kiloparsecs was $90 \times 130 \times 550$ parsecs, the first two being due to the elliptical beam, the last one mainly to the cloud velocities. Such numbers should be kept in mind in judging the extensive graphs of observational results that will be presented by Westerhout and Schmidt (papers 4 to 6). The relative success of the standard method is due to the fact that the observations give a sufficiently coarse picture of the Galaxy. In narrowing down the band-width, and even more so in narrowing down the beam-width, we may detect many more deviations from the assumptions made in the standard method than we do now. It is difficult to imagine what new reduction methods may then be needed.

In the present method the required data have been derived from the following sources.

Galactic rotation curve $\omega(R)$: inner part from observed velocities at tangent points (Kwee, Muller and Westerhout [5]); outer part and points out of plane from a model of the mass distribution (new data have been computed by Schmidt). Random cloud motions: from the observed intensities at frequencies that are forbidden on the assumption of pure circular motion and from the requirement that over-correction should be avoided. Intensity of saturated line: from highest observed intensities and estimates of optical depth at those points (absolute calibration, i.e. conversion to T_0, is not needed at this stage).

There is no guarantee that any deviations of the average local motions from circular motion will have a systematic character over the entire Galaxy. In other words, I expect that the character of such deviations will be local and erratic and that it can be studied only by long and patient observations. But I wish to mention two suggested systematic modifications. Edmondson ([6], see also paper 3) has suggested that the gas spirals out at an angle $\phi = 4°$ with circular motion. Vera Rubin [7] has made a computation based on the assumption that the gas has a motion

$V_s = 20$ km./sec. along a spiral arm on top of the circular motion. Such assumptions give rise to a slightly distorted form of the mass distribution obtained by the standard method. On studying the precise plots of the spiral arms as found in the Leiden surveys, however, we are so impressed by the comparative irregularity that it is difficult to assign much weight to Edmondson's point that the kink in the anticentre is drawn straight by his assumption. The centre direction would give a good check, as the circular motion gives strictly zero radial velocity, but the interpretation is confused by the large optical depth. Here observations of the deuterium line might be an important help. Also an accurate comparison of northern and southern observations of the rotational velocities would be helpful to determine the value of ϕ.

I wish to add in this connexion that systematic effects of over 1 km./sec. would also be observed if the nebular red-shift would not be due to recession but would be a cosmological distance effect that worked already inside the Galaxy.

6. HIGH-LATITUDE SURVEYS

The standard method has been successful in the sections of the Milky Way near the galactic plane, $|z| < 500$ parsecs or so, and farther away than 1 or 1·5 kiloparsecs. The study of the nearer regions requires high-latitude observations. Moreover, high latitude observations probably give information mostly on the gas in our immediate neighbourhood. Bok and his associates have concentrated on this subject. One problem they have studied in some detail is the association of neutral hydrogen gas with dark clouds. I shall not try to relate the details, which are reviewed in papers 7 to 10.

Helfer and Tatel ([8], see also paper 11) have detected at $l = 50°$ and $90°$, $b = 20°$ to $40°$, extensions to negative velocities up to 60 km./sec. The fact that similar extensions seem to be common at still higher latitudes, from 40° to 90°, in the measurements made at Kootwijk (see paper 3), suggests that they are not located in the next spiral arm, but are due to a more extended medium embedding the arms.

7. ABSORPTION EFFECTS

Absorption of radiation from a continuous source by the interstellar hydrogen gas at 21 cm. gives information not otherwise available. The main asset is the higher angular resolution, which is in this case determined by the point source, if the observations are correctly interpreted.

The theory is simplest if we first make the *incorrect* assumption that the gas has a constant temperature T_h and, at a given radial velocity, a distribution in the line of sight that is constant over the solid angle Ω_b of the antenna beam. Further, let us suppose that the source is characterized by a solid angle Ω_s, an optical depth τ_s, and an average surface brightness $T_s(1-e^{-\tau_s})$. Finally, let τ_1 be the optical depth of the hydrogen gas in front of the source and τ_2 the optical depth of the hydrogen behind the source, all at a certain frequency.

The antenna temperatures that might be measured by a direct radiometer are

T_{SH}: source and hydrogen,

T_S: source only, i.e. off-frequency,

T_H: hydrogen only, the so-called expected profile, which is obtained by an interpolation between off-source observations.

From the definitions we have, after a simple reduction:

$$T_H = T_h \{1 - e^{-(\tau_1+\tau_2)}\}$$

$$T_S = \frac{\Omega s}{\Omega b} T_s \{1 - e^{-\tau_s}\}$$

$$T_{SH} = T_H + e^{-\tau_1} T_S - T_h \frac{\Omega_s}{\Omega_b} e^{-\tau_1} (1 - e^{-\tau_s}) (1 - e^{-\tau_2}).$$

A comparison radiometer records the difference

$$\Delta T = T_{SH} - T_S,$$

which is positive if the radiation from the solid angle $\Omega_b - \Omega_s$ is preponderant and negative if the absorption of radiation from the source predominates. The positive part may be eliminated by subtracting the 'expected profile', thus leaving

$$\Delta T - T_H = -(1 - e^{-\tau_1}) T_S - T_h \frac{\Omega_s}{\Omega_b} e^{-\tau_1}(1 - e^{-\tau_s}) (1 - e^{-\tau_2}).$$

This difference is always negative and gives a virtually correct idea of the absorption in front of the source, for the second term can usually be neglected, e.g. if no hydrogen is behind the source, or the source is transparent, or small.

Dropping now the assumptions made, we see at once that the last formula should be still correct if there is no homogeneity of the gas across the full beam, simply because the difference is based on the solid angle Ω_s only.

However, another theorem, that followed from the assumptions, no longer holds true, namely that $T_{SH} \geqslant T_h$ whenever $T_S \geqslant T_h$. This theorem

may be derived directly from the formulae above. Its physical basis may be explained as follows. Whenever a lot of hydrogen is placed in front of the source in order to repress the source radiation, the radiation by this hydrogen itself approaches the saturation value T_h in the solid angle of the source, so contributes the small amount $T_h \Omega_s/\Omega_b$ to the antenna temperature and gives insufficient compensation. However, the equivalent hydrogen that is supposed to be in the full beam of the antenna gives sufficient compensation, bringing the antenna temperature to at least T_h. It is now clear that in an inhomogeneous gas directions and frequencies may occur in which this compensation is absent, so that the source radiation is almost fully repressed without compensation by the radiation with temperature T_h from the full beam. The observed effect then is that $T_{SH} < T_h$, or $\Delta T < T_h - T_s$. The fact that such cases seem to have been observed by the N.R.L. group gives the first definite proof of the existence of a fine structure in the gas distribution, well below the antenna resolving power.

8. EXTRAGALACTIC STUDIES

The Australian radio astronomers have made a fine study of the Magellanic clouds (Kerr, Hindman, Robinson[9]) thus adding substantially to our knowledge of these companions of our Galaxy. Beyond that, no 21-cm. radiation has been detected from any object, or group of objects. The reason is the large velocity spread giving wider line profiles than in the galactic studies. In Holland, we have been un-equipped, so far, to tackle this problem seriously as the present switching interval is only 1·08 Mc./s. and the band-width is much smaller than may be profitable for such a study.

Perhaps a rough prognosis may be useful. The Andromeda nebula, M31, assumed distance 500 kiloparsecs, is the first to be tried. The smaller angular size of M33 is not sufficiently compensated by the smaller velocities in the line of sight in this nearly face-on nebula.

The computation of a line profile for a model nebula may be made by means of the formula

$$nl = 0{\cdot}0006 \int T^* dv$$

where v is the velocity in the line of sight in km./sec., T^* is the brightness temperature, uncorrected for self-absorption, of the nebula at this velocity at a certain point in the projected nebula, l is the length of the line of sight in kiloparsecs through the nebula at that point, and n is the average number of hydrogen atoms/cm.^{3} along this line. The correction for self-