The following full text is a preprint version which may differ from the publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/75906

Please be advised that this information was generated on 2019-08-25 and may be subject to change.
Science with a lunar low-frequency array:  
From the dark ages of the Universe to nearby exoplanets

Sebastian Jester  
Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany

Heino Falcke  
Department of Astrophysics, Institute for Mathematics, Astrophysics and Particle Physics, Radboud University, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands

ASTRON, Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands

Abstract

Low-frequency radio astronomy is limited by severe ionospheric distortions below 50 MHz and complete reflection of radio waves below 10–30 MHz. Shielding of man-made interference from long-range radio broadcasts, strong natural radio emission from the Earth’s aurora, and the need for setting up a large distributed antenna array make the lunar far side a supreme location for a low-frequency radio array. A number of new scientific drivers for such an array, such as the study of the dark ages and epoch of reionization, exoplanets, and ultra-high energy cosmic rays, have emerged and need to be studied in greater detail. Here we review the scientific potential and requirements of these new scientific drivers and discuss the constraints for various lunar surface arrays. In particular we describe observability constraints imposed by the interstellar and interplanetary medium, calculate the achievable resolution, sensitivity, and confusion limit of a dipole array using general scaling laws, and apply them to various scientific questions.

Of particular interest for a lunar array are studies of the earliest phase of the universe which are not easily accessible by other means. These are the epoch of reionization at redshifts $z=6-20$, during which the very first stars and galaxies ionized most of the originally neutral intergalactic hydrogen, and the dark ages prior to that.

For example, a global 21-cm wave absorption signature from primordial hydrogen in the dark ages at $z=30-50$ could in principle be detected by a single dipole in an eternally dark crater on the moon, but foreground subtraction would be extremely difficult. Obtaining a high-quality power spectrum of density fluctuations in the epoch of reionization at $z=6-30$, providing a wealth of cosmological data, would require about $10^5-10^6$ antenna elements on the moon, which appears not unreasonable in the long term. Moreover, baryonic acoustic oscillations in the dark ages at $z=30-50$ could similarly be detected, thereby providing pristine cosmological information, e.g., on the inflationary phase of the universe.

With a large array also exoplanet magnetospheres could be detected through Jupiter-like coherent bursts. Smaller arrays of order $10^2$ antennas over $\sim 100$ km, which could already be erected robotically by a single mission with current technology and launchers, could tackle surveys of steep-spectrum large-scale radio structures from galaxy clusters and radio galaxies. Also, at very low frequencies the structure of the interstellar medium can be studied tomographically. Moreover, radio emission from neutrino interactions within the moon can potentially be used to create a neutrino detector with a volume of several cubic kilometers. An ultra-high energy cosmic ray detector with thousands of square kilometre area for cosmic ray energies $>10^{20}$ eV could in principle be realized with some hundred antennas.

In any case, pathfinder arrays are needed to test the feasibility of these experiments in the not too distant future. Lunar low-frequency arrays are thus a timely option to consider, offering the potential for significant new insights into a wide range of today’s crucial scientific topics. This would open up one of the last unexplored frequency domains in the electromagnetic spectrum.

Key words: Instrumentation: interferometers, Cosmological parameters, Neutrinos, Surveys  
PACS: 95.55.Jz, 95.80.+p, 95.85.Ry, 98.80.-k

1. Introduction

Low-frequency radio astronomy is currently experiencing an impressive revival with a number major new facilities in operation or under construction, such as the Giant Metrewave Radio
Telescope (GMRT, Swanep et al., 1991)\(^1\), Low-Frequency Array (LOFAR, Falcke et al., 2006; Roettgering et al., 2006)\(^2\), Long-Wavelength Array (LWA, Kassim et al., 2005b)\(^3\), Murchison Widefield Array (MWA, formerly known as the Milleuon Widefield Array, Morales, 2005; Morales et al., 2006)\(^4\), the 21 Centimeter Array (21CMA)\(^5\); formerly called Primeval Structure Telescope, PAST\(^6\), Peterson et al., 2005), and Precision Array to Probe the Epoch of Reionization (PAPER, Bradley et al., 2005). Eventually the Square-Kilometre Array (SKA, Schilizzi, 2004; Carilli and Rawlings, 2004)\(^7\) may extend the expected collecting area even further. Main science drivers of these telescopes are the so-called epoch of reionization (e.g., Shaver and de Bruyn, 2000; Carilli et al., 2004, and references therein), large scale surveys, transient and variable source monitoring (Fender et al., 2006), observations of our own sun (Kasper, 2006) and solar system (Zarka, 2000), as well as exoplanets (Zarka, 2007) and astrophysical and astrophysical physics (Falcke and Gorham, 2003). This rapid development together with the lively discussion of a possible return to the lunar surface by various space agencies, has raised again the long-standing question about the potential of low-frequency astronomy from space (e.g., Lazio et al., 2007).

The currently planned ground-based telescopes will provide a serious advance in radio astronomy and extend the accessible frequencies to the widest range possible from the ground. The Earth’s turbulent ionosphere gives rise to “radio seeing”, making ground-based radio observations of the sky become difficult at \(v \lesssim 100\,\text{MHz}\). At even longer wavelengths below about 10–30 MHz one encounters the ionospheric cut-off where radio waves are reflected, permitting long-distance short wave transmission around the Earth, but prohibiting observations of the sky. Observing just above the cut-off, i.e., between \(10–50\,\text{MHz}\) requires especially favourable geomagnetic and ionospheric conditions to obtain any decent images. The range below the cutoff is only readily observable from space. Hence, the dominant “low-frequency/long-wavelength” regime for which ground-based telescopes are being designed is at frequencies \(> 30\,\text{MHz}\) and wavelengths \(< 10\,\text{m}\). We will therefore hereafter refer to wavelengths above 10 m as the ultra-long wavelength (ULW) range. This wavelength-based definition is preferable to a frequency-based one to avoid confusion with the technical very low frequency (VLF) designation that extends from 3-30 kHz; the range 300 kHz–30 MHz is in engineering language named the medium and high frequency regime (MF & HF), which is clearly not appropriate in a modern astronomical context.

The best resolution achieved so far in the ULW range is on the scale of around 5 degrees, but more typically of order tens of degrees. This compares rather unfavourably to the milli-arcsecond resolution that can be routinely obtained in very long baseline interferometry (VLBI) at higher radio frequencies. Hence, the low-frequency Universe is the worst-charted part of the radio spectrum, and perhaps even of the entire electromagnetic spectrum. By today, only two kinds of maps of the sky have been made at frequencies below \(30\,\text{MHz}\). The first are maps of a part of the southern sky near the Galactic center such as those obtained by Cane and Whitham (1977), Ellis and Mendillo (1987) and Cane and Erickson (2001) from Tasmania. These have angular resolutions ranging from 5 to 30 degrees. The second kind are the maps obtained by the RAE-2 satellite (Novaco and Brown, 1978) with angular resolution of 30 degrees or worse. None of these maps show individual sources other than diffuse synchrotron emission of the Galaxy, nor do they cover the entire sky.

To improve this situation and to overcome these limitations, space-based low-frequency telescopes are required for all observations below the ionospheric cutoff (Weiler, 1987; Weiler et al., 1988; Kassim and Weiler, 1990). This is also true for a significant part of the seeing-affected frequency range above the cutoff frequency where high-resolution and high-dynamic range observations are required, such as imaging of 21-cm emission of neutral hydrogen in the very early Universe (Carilli et al., 2007).

So far, there have only been two space missions whose primary purpose was low-frequency radio astronomy: the Radio Astronomy Explorers (RAE) 1 and 2. The design and results of these missions have been reviewed by Kaiser (1987). Kaiser (1990) summarized the science goals of the RAE missions as four-fold, namely characterizing the spatial and spectral structure of Galactic radio emission, observing bursts from the Sun and Jupiter, and attempting to detect discrete Galactic and extragalactic sources. It was a surprising finding of RAE-1 that the Earth itself is a strong emitter of low-frequency bursts, the so-called Auroral Kilometric Radiation (AKR) generated by solar-wind interactions with the Earth’s magnetic field. This emission is so strong that RAE-2 was placed in a lunar instead of the originally planned terrestrial orbit to provide shielding from this natural terrestrial interference (Fig. 1 and §2.3). Of the original science goals, only the solar and Jupiter burst studies could be completed; in addition, terrestrial bursts were studied. The structure of the Galaxy’s emission was only seen at very low spatial resolution (with beam sizes between \(37' \times 61'\) and \(220' \times 160'\)) and fairly low signal-to-noise ratio. Due to the very limited angular resolution and the large power of AKR, it proved impossible to image any discrete sources directly. Thus, two of the initial scientific questions of the RAE missions remain unanswered to date.

Low-frequency radio imaging telescopes at these frequencies are nowadays realized as digitally phased arrays of thousands of simple dipole antennas that are spread over a large area. This puts a serious constraint on any realization consisting of freely-flying antennas in space and very likely favours lunar surface-based concepts in the long term. Moreover the moon provides unique shielding against disturbing radio interference from sun and Earth and within the exploration efforts of space agencies lunar missions appear to be a realistic possibility. Hence we will focus here on a lunar surface telescope concept consisting

---

\(^1\) http://www.gmrt.ncra.tifr.res.in
\(^2\) http://www.lofar.org
\(^3\) http://lwa.umn.edu
\(^4\) http://www.haystack.mit.edu/ast/arrays/mwa/
\(^5\) http://21cma.bao.ac.cn/
\(^6\) http://web.phys.cmu.edu/~past/overview.html
\(^7\) http://www.skatelescope.org
of dipoles as the most basic antenna unit. Nonetheless, most of our conclusions are general enough that they can be of use also for other kinds of implementations.

Over the past four decades several conferences and study groups formed by space agencies have produced reports on lunar radio astronomy (Weiler, 1987; Weiler et al., 1988; Kassim and Weiler, 1990; Nieuwenhuizen et al., 1992; Landecker et al., 1991, 1992; Dainty, 1992; Foing, 1994, 1996; Bély et al., 1997; Woan, 1997; Jones et al., 1998, 2000; Stone et al., 2000; Takahashi, 2003; Oberoi and Pinçon, 2005; Kassim et al., 2005a). Starting from these earlier studies we will here summarize the prospects and fundamental limitations of ULW observations and discuss their scientific potential in light of our current knowledge. Recent scientific advances in many areas and in particular in cosmology lets low-frequency astronomy shine in a new light and makes this a timely topic.

The plan of the paper is as follows. We describe observability constraints for ULW observations from the Earth and the Moon in §2, including a description of the Galactic foreground emission, and formulae for the estimation of array sizes, sensitivities etc. We give science questions to be addressed by a lunar low-frequency array in §3 and summarize our findings in §4.

2. Observability constraints for ULW observations

2.1. Earth Ionosphere

Radio waves are scattered by free electrons, such as those occurring in a space plasma or the Earth’s ionosphere. The scattering strength increases towards lower frequencies. At frequencies below the plasma frequency, which increases with electron density, radio waves cannot propagate at all. The plasma frequency of the Earth’s ionosphere is typically near 10 MHz by day and near 5 MHz by night, implying that radio observations near and below these frequencies are impossible from the Earth (the plasma frequencies can reach somewhat lower values near the Earth’s magnetic poles). At frequencies up to a few 100 MHz, the ionosphere causes celestial radio sources to suffer significantly from angular displacements, angular broadening, and intensity fluctuations (scintillation). These features are akin to seeing, the twinkling of stars visible to the eye, which is caused by density fluctuations of turbulence layers in the Earth’s atmosphere. Because of the scattering, the best available ground-based maps of the low-frequency sky have resolutions ranging from 30 degrees at 2.1 MHz to 5 degrees at 10 MHz (Ellis and Mendillo, 1987; Cane and Erickson, 2001) and very poor dynamic range.

2.2. Lunar Surface and Ionosphere

Even though the moon does not have a gaseous atmosphere, instruments carried by the Apollo 14 and Luna 19 and 22 missions showed that the moon’s surface has a photoelectron sheath, i.e., a weak ionosphere. On the day side, the inferred electron densities lay in the range 500–10,000/cm³, with plasma frequencies between 0.2 and 1 MHz (see Reasoner and Burke 1972, Benson et al. 1975 and references therein). Electron densities on the night side are expected to be much lower, since the moon’s surface potential becomes negative. However, the electron density has not been mapped...
yet in detail as function of lunar position and altitude above the surface. This is clearly a necessity, given that electron densities at the high end of the inferred ranges would mean that a lunar observatory only gains a factor of 10 in frequency during lunar day compared to what is observable below the Earth’s ionosphere. The current generation of moon missions, in particular the Japanese KAGUYA/SELENE mission, should clarify this issue.

Secondly, low-frequency radio waves can propagate into the lunar regolith to depths of 1-100 wavelengths. Subsurface discontinuities in the electrical properties of the lunar regolith could therefore lead to reflections and hence stray signals being scattered into the beam of lunar antennas. Thus, the subsurface electrical properties of the prospective sites for a lunar radio telescope need to be explored before construction of a very large array (compare Takahashi, 2003).

2.3. Man made and natural Interference

The low-frequency radio spectrum is occupied to a large fraction by terrestrial broadcasts, in particular FM and longer-wave radio and television broadcasts and military and civil communications. This radio frequency interference (RFI) is a severe problem for all kinds of ground-based radio telescopes. These signals block out cosmic signals directly at the corresponding wavelengths, but they also lead to an increased noise level for observations at other frequencies through intermodulation products that are picked up in the receiving system, through leakage, or if the affected frequencies are not filtered out completely.

At the longest wavelengths, these signals are reflected by the ionosphere. This is of course the very reason why these wavelengths are used for world-wide radio broadcasts, but it also means that a terrestrial ULW radio telescope is sensitive not just to local interference, but also to interference from all parts of the world, irrespective of location.

When the first dedicated radio astronomy satellite, RAE-1, was designed, it was expected that placing the satellite in a low orbit above the ionosphere would lead to a sufficient suppression of interference. However, it turned out that there were very strong radio bursts from the Earth itself, generated by the interaction between solar-wind particles and the Earth’s magnetic field, in particular in the frequency range 150-300 kHz. At higher frequencies, man-made and lightning interference are also still detectable above the ionosphere. Therefore, the second mission, RAE-2, was redesigned and this satellite was put into an orbit around the Moon. Figure 1 shows how well terrestrial radio interference is shielded by the Moon.

Naturally, also the sun and Jupiter are strong sources of ULW radio emission - especially during outbursts - and again here the moon can act as a perfect shield for a large fraction of the time.

Takahashi (2003) has modelled the reflection and absorption of low-frequency radio waves by the Moon, including the shielding by mountains. He finds that a large fraction of the incident energy is reflected by the lunar surface (55% at 0.5 MHz), and an effective skin depth of 2.2 km (v/MHz)^-0.81.
The distributions of electrons and magnetic fields constitute the dominant foreground (an active study. The component of the ISM that contains the free electrons responsible for the scattering is called the Warm Ionized Medium (WIM). It broadly follows the distribution of gas in the Galaxy, i.e., it is concentrated in the plane of the Galaxy. The brightness temperature of the emission rises from about 10^4 K at 30 MHz to about 2.6 x 10^7 K at around 1 MHz, and then increases less rapidly (and the specific intensity decreases) because free electrons absorb low-frequency radiation via free-free absorption (see Figure 3 and following section).

For reference, the conversion between flux density (Jansky, Jy) and brightness temperature (in Kelvin, K) is given by

\[ S_B = B \Omega \frac{T_B}{10^8 K \ deg^2} \frac{\nu}{10 MHz}^2 \]  

where \( B \) is the surface brightness, \( T_B \) is the surface brightness temperature, \( S_B \) is the flux density, \( \Omega \) is the source’s solid angle, and \( k \) is the Boltzmann constant 1.38 x 10^{-23} J/K.

Because of the absorption, the surface brightness distribution of the Galaxy varies with frequency: at high frequencies, the Galactic plane is seen in emission, with HII regions being superimposed in absorption, while the poles are more or less transparent; at frequencies around 2-3 MHz, the surface brightness is constant in all directions, so that the sky appears equally foggy everywhere; at lower frequencies, the plane is seen in absorption, so that the poles will appear warmer than the plane.

The main effect of the Galactic synchrotron emission at low frequencies is that the sky temperature is so high that it dominates the system temperature and therefore limits the sensitivity. Figure 4 shows the expected sensitivity for a lunar array as a function of the number of crossed dipoles, each with effective area

\[ A_{\text{eff}} = \frac{\lambda^2}{4}, \]  

where effective area is related to primary beam size \( \Omega \), which is also the size of the antenna’s field of view, by

\[ A_{\text{eff}} = \frac{\lambda^2}{\Omega}. \]  

We use the RMS sensitivity (1\( \sigma \)) of an array of antennas (Thompson et al., 2001, the origin of this expression is the system equivalent flux density, SEFD, which is the surface brightness corresponding to the system temperature) as given by Cohen (2004):

\[ \sigma_{\text{RMS}} = 58.9 \text{ mJy/beam} \times \frac{T_{\text{sys}}}{A_{\text{eff}} \sqrt{N(N-1)} N_{\text{IF}} t_{\text{exp}} \Delta v}. \]  

where \( A_{\text{eff}} \) is the antenna’s effective area in m², \( \Delta v \) and \( t_{\text{exp}} \) are the bandwidth in MHz and observing time in hours, respectively, over which the signal is integrated, and \( N \) is the number of array elements (for a single dipole, the term \( N(N-1) \) needs to be replaced by 1). \( T_{\text{sys}} \) is the system temperature, i.e., the sky temperature (given in Fig. 3) and approximated here as

\[ T_{\text{sky}} = \begin{cases} 16.3 \times 10^6 K \left( \frac{\nu}{2 \text{MHz}} \right)^{2.53} & \text{at } \nu > 2 \text{ MHz} \\ 16.3 \times 10^6 K \left( \frac{\nu}{2 \text{MHz}} \right)^{-0.3} & \text{at } \nu \leq 2 \text{ MHz}. \end{cases} \]  

For a dish antenna, \( N_{\text{IF}} \) is the number of intermediate frequency (IF) bands, i.e., 1 or 2 depending on whether only one or both polarizations are recorded simultaneously; for our crossed dipoles, the effective area already includes the factor of 2 for recording both polarizations simultaneously, so that \( N_{\text{IF}} = 1 \) is appropriate here.

For later reference, we can also cast eqn. 4 in terms of surface brightness sensitivity:

\[ T_{\text{RMS}} = \frac{D_{\text{max}}^2 T_{\text{sys}}}{A_{\text{eff}} \sqrt{N(N-1)} N_{\text{IF}} t_{\text{exp}} \Delta v } = \frac{f \ T_{\text{sys}}}{\sqrt{N_{\text{IF}} t_{\text{exp}} \Delta v }}, \]  

where \( D_{\text{max}} \) is the maximum baseline of the array setting the maximum angular resolution and

\[ f = \frac{A_{\text{eff}} \sqrt{N(N-1)}}{D_{\text{max}}^2} \]  

is the filling factor of the array, i.e., roughly the fraction of the array’s physical area that is filled with effective antenna area.
There is a fundamental maximum to the useful number of dipoles within a given area: an array with resolution $\theta$ arcminutes can have at most

$$N_{\text{max}} = 4.7 \times 10^7 \left( \frac{\theta}{\lambda} \right)^{-2}$$

(9)
dipoles before the filling factor exceeds unity, i.e., the dipoles start becoming mutually coupled; the maximum baseline necessary to reach this resolution is given by

$$D_{\text{max}} = 3438 \lambda \left( \frac{\theta}{\nu} \right)^{-1}$$

(10)

We consider a related issue in §2.5 below: the maximum number of useful antennas before confusion noise (rather than random noise) limits the array performance.

2.4.2. Free-free absorption

The magnitude of the free-free (or thermal bremsstrahlung) absorption coefficient $\kappa_\nu$ depends on both electron density and temperature as

$$\kappa_\nu = 1.78 g_{\text{ff}} \frac{n_e^2}{\nu^2 T_e^{3/2}} \text{ m}^{-1}$$

(11)

where $n_e$ is the density in cm$^{-3}$ and $T_e$ the temperature in K of the electron distribution, $\nu$ is the observing frequency in Hz, and $g_{\text{ff}}$ is the free-free Gaunt factor

$$g_{\text{ff}} = 10.6 + 1.6 \log T_e - 1.26 \log \nu$$

(12)

(Peterson and Webber, 2002). For the parameters of the interstellar medium $n_e \approx 0.025$ cm$^{-3}$ and $T_e \approx 7000$ K (Peterson and Webber, 2002), the distance at which the ISM optical density satisfies $\tau = 1$ is about 1.5 kiloparsec (kpc) at 1 MHz; this distance scales as the square of the frequency. For comparison, the thickness of the Galactic disc, in which the free electrons are found, is about 1 kpc, and the diameter of this disc is about 25 kpc. Thus, extragalactic observations are impossible below about 3 MHz unless there are low-density patches. Indeed, the Galactic synchrotron spectrum (Fig. 3) turns over at frequencies near 3 MHz due to absorption (this turnover is one of the ways in which the Galactic electron density and temperature are constrained observationally). Variations of the electron temperature and density in the ISM will already cause difficulties in interpreting data in the range 3-7 MHz, since variations in the surface brightness arising from fluctuations in the foreground emission and absorption have to be disentangled from structures in the sources. For some kinds of sources, this may be done by comparing the source spectra as well as their morphology, but for extragalactic synchrotron sources with similar spectra to the Galactic synchrotron emis-
tion, this task will be considerably more difficult the more extended they are.

On the other hand, the absorption enables a tomography of the ISM, since observing at different frequencies picks out absorbing structures at different distances. This is described in section 3.3.1.

2.4.3. Brightness temperature and dynamic range

Below about 100 MHz, the brightness temperature of the Galactic emission is much higher than that of any noise source in the telescope itself and therefore sets the system temperature, $T_{\text{sys}}$. Because of the structure of Galactic emission and absorption, $T_{\text{sys}}$ at any given frequency will vary strongly with direction (for example, below 30 MHz, HII regions will be seen in absorption against the Galactic plane). Erickson (1999) noted that this can lead to dynamic range and calibration problems already when observing with the VLA’s steerable antennas at 74 MHz; these difficulties are likely to be exacerbated by the use of lower-frequency dipoles with larger primary beams. The operation of LOFAR will provide experience and expertise in the calibration of an interferometer in the presence of strongly variable backgrounds. For some applications a dynamic range of $> 10^5$ is sought. Since the field of view of dipole antennas is practically the entire sky, the necessary dynamic range can be estimated by comparing the flux densities of the brightest radio sources in the sky, the supernova remnant Cas A (diameter $\approx 6\arcmin$, flux density 65 kJy) and the radio galaxy Cyg A (diameter $\approx 2\arcmin$, flux density 31.7 kJy); both flux densities are taken from Tab. 2 of Baars et al., 1977), to the desired flux limit. For extragalactic sources, the latter is the confusion limit (see eq. 22 below). At 15 MHz and with a resolution of $2\arcmin$, for example, the confusion limit is 0.14 Jy, requiring a dynamic range of $\approx 2 \times 10^5$ to image sources at the confusion limit against the glare of Cyg A.

2.4.4. Scattering in the Interplanetary and Interstellar Media (IPM/ISM)

Like all solar-system planets, the Earth is embedded in the interplanetary medium (IPM): the solar wind, an extension of the solar corona. The properties of the solar wind vary with the 11-year solar activity cycle. The solar system itself is embedded in the ISM, as already discussed above. Electron density fluctuations both in the ISM and the IPM cause scattering of radio waves, leading to the phenomena of diffractive and refractive scintillation, scatter broadening, temporal broadening of pulses, and Faraday depolarization of linearly polarized sources. The angular scattering and scintillation are akin to “seeing”. This is the blurring of optical images in ground-based telescopes caused by density fluctuations in turbulent layers of the Earth’s atmosphere and the main motivation for optical space telescopes and adaptive optics.

2.4.5. Angular scattering

At frequencies below 30 MHz, the scattering is in the “strong” regime for both media (Cohen and Cronyn, 1974), which means that there are no intensity scintillations, and the scattering broadens the intrinsic source size into a scattering angle that scales roughly as $\theta_s^2$. There is solar-wind scattering at all elongations from the sun, but the scattering is smallest in the anti-solar direction and increases for lines of sight passing closer to the sun. Similarly, the ISM is completely surrounding the Earth so that there is some scattering in all directions, but because of structures in the ISM, the scattering is stronger for lines of sight passing through the plane of the Galaxy and closer to the Galactic centre.

The expected order of magnitude of the scattering for both media can be extrapolated from measurements at higher frequencies. Empirically, the ISM broadening of the image of an extragalactic source is roughly given by

$$\theta_{\text{ISM}} \approx \frac{30'}{(\nu/\text{MHz})^{2.2} \sin b},$$

where $b$ is the Galactic latitude (Cohen and Cronyn, 1974). Estimates of the IPM scattering at 1 MHz range from a factor 3 smaller than the ISM scattering (Rickett and Coles, 2000) to a factor 5 greater (Woan, 2000). Figure 5 shows the achievable resolution with the limits imposed by IPM and ISM scattering, as function of array size, using the worst-case IPM scattering given by Woan (2000):

$$\theta_{\text{IPM}} \approx \frac{100'}{(\nu/\text{MHz})^2},$$

and a slightly different version of eq. 13 for ISM scattering, $\theta_{\text{ISM}} = 22'(\nu/\text{MHz})^{2.2}$. The best-case IPM scattering (Rickett and Coles, 2000) would be a factor of 15 smaller than given by eq. (14).

The scaling $\theta \propto \nu^2$ arises from a thin-screen approximation for the scattering medium. For a screen of thickness $L$ with overdensities of characteristic radius $a$ and excess density $\Delta N_e$, the scattering angle is given by $\theta_{\text{thin}} = r_e a^2 \Delta N_e \sqrt{(L/a)}/2\pi$, where $r_e = 2.82 \times 10^{-16}$ m is the classical electron radius. A more realistic model of the ISM/IPM accounts for their turbulent nature, yielding the $\nu^2$ scaling used above. The strength of the scattering is set by the scattering measure, a weighted line-of-sight integral of the turbulence power-spectrum coefficients (see Cordes et al., 1991 and Cordes and Lazio, 2002, who also give expressions relating the scattering measure to observables, as well as the emission measure and the pulsar dispersion measure; Cordes and Lazio, 2002 furthermore report a model for the distribution of free electrons in the Galaxy and give a computer programme that allows computation of the scattering measure, angular broadening, etc.).

2.4.6. Temporal broadening

Together with angular broadening, electron density fluctuations due to ISM and IPM turbulence also cause temporal broadening of radio pulses. The temporal broadening arises because the angular broadening allows the emission to reach the observer via multiple optical paths with different lengths. The magnitude of temporal broadening $\Delta \tau_\text{B}$ is just the geometric delay between the scattered and unscattered signal paths and therefore related to the magnitude of angular broadening $\theta_s$ by
Fig. 5. Resolution of a radio array as function of baseline. The ISM and IPM scattering limits are also shown, assuming the scattering laws given by Woan (2000). The ISM scattering size is in principle a firm limit to the resolution achievable for extragalactic observations, but could be smaller in some small part of the sky, in particular away from both the Galactic plane and the Galactic centre. The IPM scattering has not been measured well at low frequencies and could be a factor 15 smaller than given. The IPM scattering is given here for the anti-solar direction; it will increase when observing closer to the sun, and vary with the solar cycle.

\[
\Delta r_b = \frac{z'^2}{2c}
\]  

in a small-angle approximation. Here \(z'\) is the reduced distance to the scattering screen and given by \(z' = (z/z')/(z + z')\), with \(z\) the observer-screen distance and \(z'\) the screen-source distance. For IPM scattering of Galactic sources and for all extragalactic observations, \(z' \gg z\) so that \(z' \approx z\). Woan (2000) gives the following rough scalings for temporal broadening:

ISM : \(\Delta t = 6 \text{ yr (n/MHz)}^{-1.44}\)  
IPM : \(\Delta t = 0.1 \ \text{s (n/MHz)}^{-1.44}\).  

With these scaling laws, the IPM temporal broadening is acceptable at least down to 1 MHz, but the ISM broadening is still 2 hours at 10 MHz, implying that no high-time resolution work can be done for sources at extragalactic distances. For sources within the Galaxy, \(\Delta t\) increases roughly with the square of the distance to the source for fixed turbulence strength (Cordes and Lazio, 2002, eq. 9; in addition to the explicit linear dependence, there is a distance dependence in the line-of-sight integral for the scattering measure); this will be relevant for the observability of radio bursts from extrasolar planets (§3.4.2 below). However, there are now indications (P. Zarka, priv. comm.) that the pulse broadening may increase less rapidly with wavelength than expected from eq. (16); this might allow the detection of ULW pulses also from extragalactic sources after all.

2.4.7. Faraday rotation and depolarization

The plane of polarization of a linearly polarized beam undergoes the so-called Faraday rotation when it propagates through a region containing both free electrons and a magnetic field component parallel to the direction of propagation. Faraday rotation can be corrected using multi-frequency observations and the known frequency dependence of the effect. However, Faraday rotation for low-frequency observations is dominated by free electrons from the solar wind in the immediate vicinity of Earth. This makes linear polarization angles difficult to calibrate for all sources (Woan, 2000), since the RMS rotation angles can be very large and time-variable. On the positive side this may provide some measure of solar wind activity and distribution.

A much more severe effect is the cellular depolarization arising from multiple uncorrelated Faraday rotation events in the ISM. The magnitude of the effect can be found by considering the random walk of the polarization angle in a region with fluctuating electron density and magnetic fields, akin to the random walk of phase in a fluctuating electron density; assuming a constant magnetic field, the RMS phase variation is given by (Woan, 2000):

\[
\Delta \phi_{\text{RMS}} = 2.6 \times 10^{-13} \lambda^2 \Delta N_e \frac{\lambda^2 a B_{\parallel} \sqrt{L/a}}{a} \ \text{rad},
\]  

where again \(L\) is the path length through the scattering region, \(a\) is the size of an electron density fluctuation, \(\Delta N_e\) is the associated electron overdensity, and \(B_{\parallel}\) the magnetic flux density component along the line of sight (all in SI units). By inserting
into equation (18) a size $a = 10^{-12}$ parsec for individual turbulent cells of the ISM, a typical electron density of $0.3 \times 10^5 \text{ m}^{-3}$, and a typical magnetic field of 0.5 nT, the magnitude of the Faraday rotation at 1 MHz is about 1 radian in one such cell. The total RMS Faraday rotation of low-frequency radiation travelling through the ISM is therefore very large, of order 1000 radians per parsec, and therefore leads on average to depolarization. Thus, no linear polarization is observable from sources beyond our own solar system at these very low frequencies.

Circular polarization is not affected, but circularly polarized sources are extremely rare at higher frequencies, though cyclotron maser bursts have been observed from certain types of stars (e.g. Hallinan et al., 2007) and is being found in an increasing number of active galactic nuclei (Gabuzda et al., 2008, e.g.). At lower frequencies, cyclotron maser bursts are expected to be visible from extrasolar planets (see §3.4.2). Also conversion from linear to circular emission in a dense magnetized plasma (see Beckert and Falcke, 2002, and references therein for an explanation) is another process that can generate some level of circularly polarized emission across a wide frequency range. Hence, circular polarization capabilities — which can be measured by cross-correlating linearly polarized feeds — should be a more promising polarization signal to look for at low frequencies.

### 2.5. Confusion and duration of a confusion limited all sky survey

The first low-frequency radio surveys suffered badly from confusion: the presence of unresolved sources with individual flux densities below the detection limit leads to a constant floor in the noise level that is reached after a certain observing time. This can also lead to source positions being substantially in error (see, e.g., Kellermann, 2003). For low-frequency radio astronomy, a greater level of confusion may arise from fluctuations in the Galactic foreground emission rather than the presence of unresolved discrete extragalactic background sources.

Among others, Di Matteo et al. (2004) have considered the impact of these fluctuations on the observability of the redshifted 21-cm power spectrum at frequencies near 100 MHz. A detailed simulation of the Galactic foreground emission based on an extrapolation of results from WMAP and LOFAR will provide useful guidance on the expected level of foreground fluctuations which need to be removed in order to detect extragalactic sources. For the case of observing redshifted line emission, spectral information will help in isolating the signal from the foreground, and transient detection will be affected predominantly by the bright foreground emission; however, the observability of extragalactic continuum sources in the presence of foreground fluctuations needs to be assessed in detail with suitable simulations.

To estimate the confusion limit for extragalactic observations, which can be done down to 3 MHz, we use the approach and formulas of Cohen (2004), who estimates the anticipated confusion limit due to background sources by extrapolating from extragalactic source counts and the observed confusion limit at 74 MHz. As a conservative estimate based on comparing the RMS noise level of 74 MHz observations in the VLA’s C and D arrays, Cohen (2004) requires a source density of less than one source per $m = 12.9$ beam areas for a survey to remain confusion-free. The beam area $\Omega_b$ of a Gaussian elliptical antenna beam with major and minor axes $\sigma_x$ and $\sigma_y$, respectively, is given by

$$\Omega_b = 2\pi \sigma_x \sigma_y$$

where the approximation holds for a circular Gaussian beam ($\sigma_x = \sigma_y = \sigma$) of full-width at half-maximum (FWHM) $\theta = \sqrt{8\ln 2} \sigma \approx 2.35 \sigma$. The confusion limit $S_{\text{conf}}$ is then defined by

$$1.13 \, \theta^2 \, m \, N_\gamma (S_{\text{conf}}) = 1,$$

where $N_\gamma (S)$ is the number of sources with flux densities greater than $S$ (equation 1 from Cohen, 2004). In the VLA Low-Frequency Sky Survey (VLSS) that is conducted at 74 MHz, the source counts at the faint end are given by

$$N_\gamma (S) = 1.14 \text{ deg}^{-2} \left( \frac{S}{10 \text{ mJy}} \right)^{-\beta} \left( \frac{\nu}{10 \text{ MHz}} \right)^{-0.7}.$$

We ignore here the fact that many radio sources start showing a low-frequency turnover due to synchrotron self-absorption, and hence we likely overestimate the number of source and thus in turn also the confusion limit. With this assumption, the confusion limit in Jansky is given as function of resolution $\theta$ by

$$S_{\text{conf}}(\theta, \nu) = \left( 12.9 \times 1.14 \times 1.13 \left( \theta/1' \right)^{1.54} \left( \frac{\nu}{74 \text{ MHz}} \right)^{0.7} \right).$$

The left panel of Figure 6 shows the confusion limit from equation (22) as function of maximum baseline and frequency (see Fig. 5 for the resolution as function of those quantities).

Since the confusion limit is a lower limit to the achievable noise, it implies an upper limit to the useful collecting area of an array — adding more antennas only decreases the time in which the confusion limit is reached, but not the overall array sensitivity. To compute the maximum useful number of antennas, we begin with the expression for the RMS sensitivity of an array of dipoles, eq. 4. As before, we assume that each array element is a pair of crossed dipoles with effective area given by eq. (2). Setting the confusion limit from eqn. 22 equal to the
1-σ RMS from eqn. 4 then yields the maximum useful number of antennas for a given integration time, which is plotted in the right-hand panel of Fig. 6 for an integration time of 24 h.

Perley (2001) gives an expression for the "survey equation", the time needed to survey a solid angle of 1 sr to a given sensitivity. This is simply the time necessary for an array to reach the desired sensitivity (via eq. 4), divided by the field of view of each array element, i.e., the primary beam area. For the case of an array of \( N \) crossed dipoles, with effective area of each element given by eq. 2, the time to survey one steradian to a 1-σ point-source flux density \( s \) is given by

\[
\mathcal{t}_{\text{survey}} = \frac{16}{3} \left( \frac{k T_{\text{sys}}}{s^2} \right)^2 \left( \frac{\Delta v}{100 \text{ kHz}} \right)^2 \left( \frac{\Delta v}{\nu} \right)^{-1} \left( \frac{\nu}{10^7 \text{ K}} \right)^{-1} \left( \frac{\nu}{10^2 \text{ MHz}} \right)^4.
\]

Inserting the high-frequency dependence of the sky temperature from eq. (5), which dominates the system temperature, choosing the confusion limit from eq. (22) as limiting flux density, and including a telescope efficiency \( \eta \), this expression can be approximated as

\[
\mathcal{t}_{\text{survey}} \approx \frac{44 d}{N^2} \left( \frac{s}{\text{Jy}} \right)^2 \left( \frac{T_{\text{sys}}}{10^7 \text{ K}} \right)^2 \left( \frac{\Delta v}{100 \text{ kHz}} \right)^{-1} \left( \frac{\nu}{\text{MHz}} \right)^4.
\]

Equation (24) gives an order-of-magnitude estimate. In practice, details like \((u, v)\) coverage will change the actual values. E.g., an array at a lunar pole will have more and more restricted \((u, v)\) coverage for sources closer and closer to the lunar celestial equator, losing the ability to resolve sources from each other in the direction perpendicular to the equator.

As just mentioned, the presence of Galactic foreground emission adds a new confusion component at frequencies below about 30 MHz. The accuracy with which extragalactic source observations can be done depends on the smallest-scale structures that are present in the foreground emission. Observations of distant radio galaxies will be hampered by the fact that both, they and the Galactic foreground, produce synchrotron emission, so that the spectral signature cannot be used for distinguishing foreground and extragalactic emission, as is possible for 21 cm spectral line emission. As noted by Graham Woan, “There are clearly many contributions to the apparent flux density of any single source, and much of the interpretation from discrete sources will therefore be statistical in nature” (Woan, 2000). However, at least at higher frequencies most of power
in the foreground is in the diffuse component and hence higher resolution clearly makes imaging more feasibly for any kind of survey work.

2.6. Stability and Calibratability

The biggest challenge for a low-frequency phased array is the calibration. The main issue for terrestrial arrays is calibrating ionospheric phase fluctuations. Due to the low-density atmosphere and ionosphere of the moon – especially during the night – we do not consider this a major problem for a lunar array. This leaves calibration of the instrumental beam shape and band pass (frequency dependent gain of the antennas and electronics) as the main concerns. For a phased array consisting of simple dipoles where all beam forming and imaging happens digitally, the overall beam pattern can in principle be calculated in a straightforward manner. However, for terrestrial applications relative gain changes in the analog parts of the dipole antennas need to be traced and calibrated out. For the low-band antennas of LOFAR the main antenna gain variations are expected to come from temperature fluctuations (affecting electronics and damping in the wires), rain (changing the electrical properties of ground and wires), and wind (leading to vibrations of the antennas). Actual variations of prototype hardware in the field at 40-80 MHz have been found to be of order 20% (Nehls et al., 2007). Since neither rain nor wind are to be expected on the moon, temperature variations will be the largest source of gain variations. Lunar surface temperatures can vary from -233°C to 132°C, well in excess of terrestrial variations. Fortunately this change in temperature happens over the course of a lunar day, i.e., 4 weeks and should be relatively smooth and predictable over the array. Also, there is no variable cloud cover that can lead to random solar irradiation. The most stable conditions on the moon would be achieved in an eternally dark spot in a polar crater, where one expects very little gain variations at all.

The final source of uncertainty in the overall calibration would be the relative location of the antennas (if not known from the deployment procedure). However, interferometric self-calibration procedures can solve for antenna positions with high accuracy, especially if the sky is dominated by a single (known) point source. This is the case during Jupiter and solar bursts and has been used effectively already to calibrate LOFAR prototype stations (such as LOPES; Falcke et al., 2005).

Hence, in contrast to arrays on Earth, calibratability does not seem to be the main driver for design and layout of the array, as long as a proper temperature control of the dipoles is achieved.

3. Science requirements

This section describes science questions to be addressed by a lunar array. As far as possible, the science requirements (resolution, sensitivity) have been translated into requirements on the array parameters (maximum baselines, number of antennas). However, there are additional requirements with regards to the two-dimensional distribution of antennas, the properties of the correlator, and the methods used to calibrate the interferometric data and process them into images of the sky. These have
not been considered in detail and are likely to modify some of the array parameters presented here. Some of these additional aspects have been discussed by Woan (1997) and Jones et al. (1998), and other previous studies for lunar or space-based low-frequency arrays (see citations at end of section 1).

3.1. Cosmology with HI line emission

A fundamental question of current cosmological research is on the nature of structure formation in the Universe (for a review, see Ciardi and Ferrara, 2005): how is the observed distribution of visible matter created from the initial conditions just after the big bang, when matter and radiation were distributed extremely smoothly, with density variations of just one part in 100,000? The CMB radiation was emitted at \( z \approx 1200 \), about 400,000 years after the Big Bang, when the Universe had cooled off sufficiently for electrons and protons to recombine into neutral hydrogen atoms (Epoch of Recombination), allowing the background radiation to move freely without being scattered by the electrons and protons. At the same time, however, the Universe became opaque to visible light, because neutral hydrogen atoms absorb visible and infrared photons and re-emit them in random directions. Moreover, there were no sources of light in the Universe yet: the hydrogen and helium that were created in the big bang first had to cool in order to be able to clump together and form stars and galaxies. Hence, this era is called the “cosmic dark ages” in the redshift range \( z = 30 – 1000 \) (Rees, 1999).

Things only changed after the first stars, galaxies and active black holes had formed and emitted enough UV and X-ray photons to reionize the neutral hydrogen, allowing all radiation to pass freely. The time when this happened is called the Epoch of Reionization (EoR) and is believed to have occurred around \( z \sim 11 \), about 400 million years after the Big Bang, though it is at present not known whether the reionization happened more or less instantaneously, similar to a global phase transition, or was more or less spread out in time, depending on local conditions.

Throughout all these epochs hydrogen played a major role, emitting or absorbing the well-known 21-cm (1.4 GHz) line due to the spin flip of the electron. This emission is redshifted by a factor 10–1000 due to the cosmological expansion and ends up in the frequency range from 140–1.4 MHz. When the hydrogen spin temperature is not coupled perfectly to the radiation temperature of the cosmic background radiation, the hyperfine populations of neutral hydrogen are in equilibrium with the background radiation, and \( \Delta T_{21} = T_s - T_R(z) = 0 \). As the first Lyman-a photons are absorbed or emitted the well-known 21-cm line due to the onset of reionization, the so-called global EoR signal. The signal is a characteristic variation of 21-cm brightness temperature \( T_{21}(z) \) with redshift, which arises at follows. Before reionization, the hyperfine populations of neutral hydrogen are in equilibrium with the background radiation, and \( \Delta T_{21} = T_s - T_R(z) = 0 \). As the first Lyman-a photons are generated, \( T_s \) becomes decoupled from the background radiation temperature and couples to the gas temperature, either by collisions or via Wouthuysen-Field scattering. Since the gas temperature is less than the CMB temperature and therefore \( \Delta T_{21} \) is negative, the 21-cm line is seen in absorption. As the gas heats up, \( \Delta T_{21} \) increases and reaches values greater than 0. The heating increases until reionization has proceeded far enough that only very few neutral hydrogen atoms are left, thus removing the source of the 21-cm emission and resetting \( \Delta T_{21} \) to its initial value of 0. The observed wavelength at which this signal appears directly reveals when reionization occurred, and the redshift evolution of \( \Delta T_{21}(z) \) encodes the detailed physics of reionization.

Because of the homogeneity and isotropy of the Universe, this transition happened roughly simultaneously in all parts of the Universe. Therefore, the same signal is observed independently of the direction one is looking in, and a single simple dipole antenna would be sufficient equipment to pick up the signal, which has an amplitude of a few to a few tens of milliKelvin (Gnedin and Shaver, 2004, Fig. 1). Thus, although simple in principle, detection of the signal is difficult against the Galactic foreground. As reionization is believed to occur in the redshift interval \( 6 < z < 20 \), the signal will appear at frequencies between 70 and 200 MHz. As discussed by Gnedin and Shaver (2004), the main problem here is not the sensitivity as such. The sky temperatures in this frequency range are 100–2000 K, with lower temperatures at higher frequencies (eq. 5). As Fig. 8 shows, the necessary observing times to reach 10-sigma detections of a 1 mK signal from \( z = 15 \) with a single dipole are just a few months. Hence, the ability to distinguish the global EoR signal from the effect of the foreground variation imposes a further
Fig. 8. Integration time necessary to reach an RMS noise level of 1 mK in the redshifted 21-cm line for observing the global Epoch of Reionization signal (of order of a few to a few tens of mK; Gnedin and Shaver, 2004, Fig. 1) with a single dipole against the Galactic foreground emission, as function of redshift \( z = 5 \) corresponds to \( v = 230 \) MHz, \( z = 30 \) to 45 MHz) The integration time is obtained from eqns. 5 and 7 with \( f = 1 \). A constant fractional bandwidth of 5% was assumed. To reach fainter noise levels the required integration time scales as \( t_{\text{exp}} \propto 1/\sigma^2 \). The redshift range shown includes the entire likely range for the origin of the global EoR signal; redshifted 21-cm emission from higher redshifts is considered separately below.

observability constraint. The evolution of \( \Delta T_{21}(z) \) corresponds to a smooth change of the spectral index as function of frequency, which needs to be distinguished from similarly smooth variations in the Galactic foreground emission. Its detection is – so far unsuccessfully – being tried from the ground and requires very high gain stability and absence of RFI: the Australia Telescope National Facility has an ongoing development effort “Cosmological Reionization Experiment” (CoRE) that attempts a global EoR measurement (Chippendale et al., 2005); Bowman et al. (2008) have just reported first results from the EDGES experiment\(^{10}\), determining an upper limit of 250 mK to the reionization signal, and stressing the need for systematic errors being limited to less than 3 mK.

A single antenna with an instantaneous frequency range 10-200 MHz, in an “eternally dark” lunar crater at the pole, kept at a stable temperature, and shielded from any solar and terrestrial radiation, would clearly offer the very best set-up for such an experiment. That would reduce most instrumental and RFI effects to the bare minimum and only leave the cosmic foreground as the main issue. Using several independent dipoles, without any beam-forming, would increase sensitivity and reliability against systematic errors even further. With enough signal paths the same antennas could then also be used in a beam-forming mode. Systematics and cosmic variance could be further checked by using both lunar poles which lets one observe two independent patches of the sky. A lunar array could also collect the global EoR signal simultaneously with other observations by adding signals from different antennas incoherently. This only requires including a dedicated signal processing chain parallel to the correlator. Together this should deliver an ultra-precise global radio spectrum of the low-frequency sky to identify the brightness temperature change in the the 21 cm line caused by reionization. This would answer the question: When did the global transition between a predominantly neutral and a predominantly ionized Universe happen?

3.1.2. The global 21-cm signal from the Dark Ages beyond reionization

In the existing literature, the focus for 21-cm cosmology has been put on the global signal from the Epoch of Reionization on the one hand, and on the power spectrum of 21-cm fluctuations from the EoR and higher redshifts on the other (see following sections). However, there is also a global signature of the dark ages from the spin temperature evolution at redshifts beyond reionization and up to \( z \approx 200 \) (e.g., Ciardi and Madau, 2003; Ciardi and Salvaterra, 2007). The expected signal is a surface brightness decrement with a peak amplitude of around 40 mK at \( z \approx 90 \) and approaching 0 at \( z \approx 20 \) and \( z \approx 400 \). We have calculated the exposure time necessary to observe this global Dark-Ages signal with a single dipole against the sky background, using the model of Ciardi and Salvaterra (2007). The

\(^{10}\)http://www.haystack.mit.edu/ast/arrays/Edges/
Fig. 9. A simulated global 21-cm HI signal from the dark ages after 1 year of integration time with a single sky-noise limited dipole as a function of frequency (based on the calculations by Ciardi and Salvaterra, 2007, also compare Fig. 10). Here a fixed bandwidth of 1.5 MHz is used. We assume that the observed brightness temperature, $T_{\text{obs}}$, is just the sum of the dark ages signal and a foreground signal, $T_{\text{for}}$, which is a perfect power law. The signal at 30 MHz originates from $z = 46$.

Fig. 10. Global 21-cm signal from HI in the dark ages. The plot shows the observing time needed to reach a 5-$\sigma$ detection at 5% fractional bandwidth with a single dipole, as function of redshift. If $N$ dipoles are added incoherently, then the observing time is reduced by a factor $\sqrt{N}$. 
result is shown in Fig. 9, with the exposure time necessary to reach a 5-σ detection for a signal peaking at different redshifts shown in Fig. 10. A single dipole can provide a 5-σ detection at \( z = 42 \) in 137 days. A caveat here is that the presence of the reionizing Lyman-α photons may be tightly correlated with high densities also beyond the very earliest stages of reionization. In this case, the final spin temperature may be much higher than the originally predicted diffuse hydrogen temperature, i.e., the spin temperature is closer to the CMB temperature, decreasing the absorption dip and weakening the absorption signature.

The simulated line detection with Gaussian noise in Fig. 9 shows that the signal is only \( 10^{-6} \) of the foreground signal. Moreover, a detection has to rely on the foreground to be a perfect power law. Hence, the foreground subtraction is the biggest uncertainty together with the need for an almost perfect band-pass calibration of the antenna. It is likely that the actual foreground signal will deviate from a simple power-law. This makes a detection almost impossible if there are spectral foreground fluctuations of the same magnitude as the dark ages signal. Nonetheless, further studies need to show on which level spectral foreground fluctuations are actually present. As in the global EoR case, the global dark-ages 21-cm signal is isotropic, so that a single (well-calibrated) dipole suffices to carry out this measurement and detect the signal from Fig. 9 in frequency space.

### 3.1.3. EoR tomography

At the other extreme of the spectrum of potential cosmology experiments would be an attempt to image the hydrogen fluctuations directly. The detailed history of reionization can be traced by observing the two-dimensional structure of neutral and reionized gas around luminous objects, and by stepping through the different frequencies corresponding to different emission redshifts, a tomographic map of reionization can be created. The characteristic scale of the fluctuations is expected at the arcminute scale (Furlanetto et al., 2004). This adds arcminute-scale spatial resolution over the frequency range 50–140 MHz (corresponding to redshifts \( 10 < z < 20 \)) to the requirements, as well as the ability to detect cosmological milli-Kelvin brightness fluctuations over the scale of one resolution element. The same hardware would allow to image hydrogen fluctuations directly also during the preceding dark ages, at \( 30 < z < 100 \), if the frequency range is extended down to 14 MHz and the baselines are increased correspondingly to achieve the desired resolution also at these longer wavelengths.

The purity requirements for EoR likely mandate a lunar far-side location to ensure absence of solar, terrestrial auroral, and man-made interference. A near-polar region could be considered since this will provide a near-constant view of the sky, allowing long integrations. If the site is closer than \(-6.5^\circ\) to the pole, lunar libration may expose the site to the Earth and make EoR observations impossible for some fraction of the year.

The expression for the exposure time necessary to reach a given surface brightness sensitivity \( T_{\text{RMS}} \) is particularly simple when expressed as function of the system temperature (which is a function of observing frequency, i.e., redshift, via \( \nu_{21}(\alpha) = 1.4 \text{GHz}/[1 + \alpha] \)) and given by solving eq. (7) for the exposure time:

\[
\tau_{\text{exp}} = \frac{1}{\Delta \nu} \left( \frac{T_{\text{sky}}(\nu_{21}(\alpha))}{T_{\text{RMS}}} \right)^2
\]

\[
= \frac{2.1 \times 10^{12} \text{yr}}{N(N-1)} \left( b \right)^{-1} \left( \frac{T_{\text{RMS}}}{1 \text{mK}} \right)^{-2} \left( \frac{\theta}{1^\prime} \right)^{-4} (1 + z)^{6.06}
\]

(25)

where \( \nu \) is the usual filling factor (eq. 8) and \( b = \Delta \nu/\nu \) is the fractional bandwidth. Here, we have used the fact that the sky temperature happens to take on the simple form

\[
T_{\text{sky}}(\nu_{21}(\alpha)) = (1 + \alpha)^{2.53} \times 1 \text{K}
\]

(26)

using the approximation to \( T_{\text{sky}}(\nu) \) from eq. (5) for the observing frequencies \( \nu > 2 \text{ MHz} \) considered here, and substituting \( \nu_{21}(\alpha) = 1.4 \text{ GHz}/(1 + \alpha) \) (see also upper panel of Fig. 12).

The exposure time from eq. (25) is plotted against target redshift in Fig. 11, assuming a scaled array in which the maximum baseline increases with wavelength to keep the resolution constant at \( \theta = 1^\prime \). This implies a filling factor \( f = 2.1 \times 10^{-8} \sqrt{N(N-1)} \); conversely, the number of dipoles required for a given filling factor at a given resolution is

\[
N_{\text{dipoles}} \approx \frac{47 \times 10^9}{f} \left( \frac{\theta}{1^\prime} \right)^{-2}
\]

(27)

for \( N_{\text{dipoles}} \gg 1 \). The corresponding effective area is \( N_{\text{dipoles}} \times A^2/4 \), i.e.,

\[
A_{\text{tot}} = 10.5 \text{km}^2 \ N_{\text{dipoles}} \left( \frac{\lambda}{30 \text{ m}} \right)^2 \left( \frac{\theta}{1^\prime} \right)^{-2}
\]

(28)

where 1 MHz/\( \nu \) can be substituted for the \( \lambda/30 \text{ m} \) term.

Again, the exposure time is driven by the large brightness of the Galactic synchrotron emission, which sets the system temperature. The exposure time can only be decreased by increasing the filling factor, i.e., the number of dipoles; since the filling factor is limited to \( f \leq 1 \), setting \( f = 1 \) in eq. (25) yields a firm lower limit to the necessary exposure time for a given \( T_{\text{RMS}} \). In principle, if enough correlator capacity is available, the imaging could be performed over a significant fraction of the entire visible sky within that time scale, since dipole antennas have a very large instantaneous field-of-view (compare the discussion of survey speed in §2.5).

It is evident that an array with a substantial number of dipoles is required to reach the sensitivity necessary for 21-cm tomographic imaging. For observing times of order one year for imaging around \( z = 20 \) (at 66 MHz or 4.4 m), of order 102 dipoles would be required, with a total effective area of 50 km2 (also see Fig. 4). The strongest factor here is the spatial resolution which enters with the fourth power in the integration time (eq. 25) and squared in the number of dipoles needed (eq. 27). Hence, a 10' array observing 21-cm emission from \( z = 20 \) (Zahn et al., 2007) would require only about 103 dipoles (i.e., 0.5 km2) at the same sensitivity limit. Higher redshifts – in the pre-EoR area – become increasingly more difficult even though the collecting area per dipole increases with \( \lambda^2 \) (i.e., \( 1 + \alpha^2 \)) because the sky temperature increases rapidly with
longer wavelengths. Fig. 4 and eq. (25) imply that at the highest redshifts the number of dipoles needed would require a filling factor larger than unity for one year of integration time and 1′ resolution. This is of course impossible and can only be remedied by using antennas with a higher gain than the crossed dipoles assumed here. For example, log-periodic or Yagi antennas have a higher gain (i.e., effective area) per antenna, at the cost of a smaller field-of-view and, worse, increased complexity compared to simple crossed dipoles. As Fig. 11 shows, a filling factor \( f = 1 \) is already needed to achieve exposure times of order 1 year for tomography at \( z \approx 15 \); since the exposure time at fixed resolution and filling factor grows roughly as \( (1 + z)^6 \), substantially higher redshifts can only be reached with very high-gain antennas or very long observing times.

Confusion noise poses further constraints: 21-cm emission from \( z = 15 \) is observed at 87.5 MHz, where the confusion noise at 1′ resolution is about 14 mJy, which corresponds to a brightness temperature of 710 K. This is much more than the desired signal of 1 mK. To resolve the confusing background and reduce it to the 1 mK level, which is prerequisite for allowing its subtraction, it would be necessary to improve the resolution to the 0′.01 level, which would require baselines larger than the diameter of the moon (3476 km). Instead, again statistical techniques have to be employed to extract the signal from the confusion noise via different structures in frequency space (Di Matteo et al., 2004).

In summary, one can conclude that a lunar EoR imaging experiment needs to be very substantial in size to penetrate into the \( z > 20 \) epoch, but is not unthinkable in the long run. This crucially depends in the expected angular size scale of the signal, which will hopefully become more clear with advances in ground based experiments. Accurate subtraction of the sources of extragalactic confusion noise is also a necessity and may prove difficult.

3.1.4. Power spectrum of the 21cm-transition at \( z = 30–50 \)

Loeb and Zaldarriaga (2004) have pointed out that the fluctuation power spectrum of the redshifted 21-cm line carries a wealth of information about the matter power spectrum. The CMB radiation itself that is observed today carries information about cosmological parameters mainly at the largest angular scales from 0°2–90° (multipole scale \( l \) from 2–1000, and limited in principle by Silk damping to \( l < 3000 \)). By contrast, the angular power spectrum in the redshifted 21-cm line carries cosmological information at much smaller angular scales, \( l > 10^4 \), corresponding to angular scales of 1′ or less. In addition, observations of redshifted 21-cm emission from different redshifts in the range 30–50 yield independent samples of the cosmological parameters, while the CMB information suffers from cosmic variance. Therefore, observations of the 21-cm power spectrum from this redshift interval yield many orders of
magnitude more information about density fluctuations in the early Universe than direct observations of the CMB power spectrum. This makes the redshifted 21-cm line a powerful tool to constrain all model parameters necessary to describe the Universe, e.g., the slope and curvature of the initial density fluctuation spectrum containing information from the inflationary phase, the mass of warm dark matter particles, and the fraction of matter density contributed by neutrinos.

In particular, angular scales \( l > 5000 \) carry the greatest weight in discriminating between different values for the cosmological parameters, since it is at these wavenumbers where the predicted power spectra for different sets of cosmological parameters differ most (Loeb and Zaldarriaga, 2004, Fig. 3). A low-frequency measurement of the 21-cm power spectrum at these angular scales and in this redshift interval would therefore be an ultimate cosmological experiment.

Baryon acoustic oscillations (BAOs) are a major tool of modern cosmology, producing a feature that is directly visible in an observed power spectrum. They are imprinted on the 21-cm angular power spectrum in the same way as on the CMB (de Bernardis et al., 2000) and galaxy power spectra (Eisenstein et al., 2005; Cole et al., 2005; Percival et al., 2007). The BAO signal is present on spatial scales of 20–400 Mpc (comoving; Barkana and Loeb, 2005, Fig. 4). In the 21-cm angular power spectrum, it is predicted to appear at wavenumbers of order \( l \approx 90–4000 \) and with oscillation amplitudes of order or milliKelvin or smaller (see, e.g., Lewis and Challinor, 2007, Fig. 8 and Mao and Wu, 2008, Fig. 4).

The distinct advantage of observing the 21-cm power spectrum from the Dark Ages, before the Epoch of Reionization, is that the astrophysics of reionization is complex, non-linear and not well understood, but these complexities do not affect the Dark-Ages power spectra, which therefore provide a much cleaner signal to observe and to interpret. For example, the non-Gaussianity of the fluctuations can be probed (Cooray et al., 2008) and may reveal information on the inflationary phase of the universe.

The required observing frequency is 1.4 GHz/(1 + \( z \)); see the upper panel of Figure 12. As noted in section 2.4.5, angular scatter broadening in the ISM and IPM limits the highest observable angular scale as a function of frequency. The limiting angular scale increases proportionally to \( \lambda^2 \) (see Fig. 5 for the maximum possible resolution and the baseline necessary for reaching this resolution). The lower panel of Figure 12 shows the number of independent samples (data points) that can be collected up to the IPM scattering limit (\( N_{\text{IPM}} \) from Loeb and Zaldarriaga, 2004, p. 4), assuming a bandwidth of 5% of the observing frequency. Since the number of samples scales as \( N_{\text{IPM}} \propto l_{\text{max}}^2 \), which in turn scales as \( (1 + z)^{-2} \), the number of samples up to the scattering limit scales as \( (1 + z)^{-6} \).

Therefore, observing fluctuations originating from \( z = 45 \) only contributes \( 10\% \) of the information accessible at \( z = 30 \), and observing \( z = 70 \) only contributes an additional \( 1\% \) compared to \( z = 30 \). Thus, as far as the number of samples is concerned, the presence of angular broadening in the ISM and IPM makes extending this experiment to lower frequencies progressively more inefficient. However, the detailed strength of the signal as a function of redshift cannot be predicted in advance, and there may be highly relevant deviations from current theory. This is particularly important since, as Loeb and Zaldarriaga (2004) point out, this phase of the cosmological evolution is a fairly clean laboratory, hardly affected by any "dirty" astrophysics of the reionization process. Measuring the evolution of structure formation with redshift between \( z = 10 - 100 \) may therefore be one of the most intriguing diagnostics of the history of the Universe. For these reasons, it will still be desirable to cover a broad redshift, i.e. frequency, range.

Sensitivity poses another constraint. The angular power spectrum has the strongest total signal at \( z = 55 \) (25 MHz), dropping off roughly symmetrically to \( z = 25 \) and \( z = 170 \) (Loeb and Zaldarriaga, 2004, Fig. 2, e.g.). The signal strength is expected to be at the mK scale, while the system temperature and hence the noise are determined by the sky temperature (see Figure 3). Figure 13 shows the observing time necessary to reach an RMS level of 1 mK in the power spectrum with an array of dipoles as given by Loeb and Zaldarriaga (2004, eq. 11). To achieve an angular resolution of 1’ in reasonable integration times of order one year, about \( 10^{5.5} \) dipoles are needed for \( z = 30 \) (\( A_{\text{eff}} = 3.5 \text{ km}^2 \)) and \( 10^6 \) dipoles (\( A_{\text{eff}} = 30 \text{ km}^2 \)) at \( z = 50 \). Again, this number depends very critically on the angular resolution required and would reduce to \( N \approx 10^{3.5} \text{ and } N \approx 10^4 \), respectively, for angular scales of 10’ if the signal were on the mK level. Distinguishing between different cosmological models at high confidence levels or detecting BAOs directly will probably require sensitivities down to the 0.01 mK level, but the latter may already be feasible with degree-scale resolution, probing multipole numbers \( l \approx 100 \) (Barkana and Loeb, 2005).

Again, accurate subtraction of spectral and noise power fore- and backgrounds is essential. Determining the exact array parameters that will be necessary requires more detailed guidance from further cosmological observations and simulations. Numerous novel approaches to extracting cosmological information from 21-cm power spectra are currently being proposed to overcome the observational challenges (see, e.g., Morales et al., 2006; Barkana and Loeb, 2008; Visbal et al., 2008; Cooray et al., 2008).

3.2. Extragalactic surveys: radio galaxies, clusters and feedback

At low radio frequencies, the brightest sources of the extragalactic sky are the synchrotron-emitting radio lobes of giant radio galaxies and quasars, such as Cygnus A. In fact, the 178 MHz 3C survey (now used in its twice-revised form 3CRR, Laing et al., 1983) was the first successful survey of discrete sources, and 3C sources are still studied extensively today. There are a number of questions that are currently of particular relevance to which the study of radio galaxies can provide decisive insights.
Fig. 12. **Upper panel:** observed frequency of redshifted 21cm emission (left-hand axis), and baseline required to reach 2' resolution at that frequency (right-hand axis), both as function of redshift $z$. Ground-based observations are exceedingly difficult below 30 MHz, corresponding to $z = 47$. **Lower panel:** Number of independent power-spectrum samples ($N_{21\text{cm}}$ from Loeb and Zaldarriaga, 2004, p. 4) obtainable down to the ISM scattering limit towards the Galactic poles, also as function of redshift.

Fig. 13. Number of antennas necessary to achieve a 5-$\sigma$ detection of fluctuations in the 21cm power spectrum in one year at multipole numbers up to $l = 5000$ (corresponding to 2' angular resolution) or $l = 500$ (20') at the 1 mK level, as function of redshift (as given by eq. 11 from Loeb and Zaldarriaga, 2004). Except for the different exposure time prescription, this figure is constructed in the same way as Fig. 11, so that the same scalings apply, i.e., $N_{\text{Ant}} \sqrt{t_{\text{exp}} S^{-1} (1 + z)^3} = \text{const}$ (where $S$ is the desired sensitivity level), assuming a filling factor independent of frequency.
3.2.1. Electron energy distributions and jet power in radio galaxies

The low-frequency emission of radio galaxies and quasars is generated predominantly in the “lobes”. These are structures with scales of tens to hundreds of kpc (Miley, 1980; Begelman et al., 1984) that are inflated by pairs of relativistic jets emerging from near the accretion disk around a black hole in the nucleus of the host galaxy.

Jets and lobes of radio galaxies emit synchrotron radiation, which has the property that the lowest-frequency emission is generated by the lowest-energy electrons. Since electron energy distributions typically are of the form \(N(E) dE \propto E^{-\alpha} dE\) with \(p \approx 2.4\) to yield the observed spectra \(\nu^{-0.7}\), low-energy electrons are the most numerous and dominate the power output. On the other hand, as discussed below, higher-energy electrons have shorter cooling time scales, leading to a steepening of the radio spectrum – the so-called spectral aging (Alexander and Leahy, 1987; Blundell and Rawlings, 2001). Hence, older parts of a source tend to have steeper radio spectra. Low-frequency observations thus probe the oldest structures in a galaxy and can be used to constrain the age of a source.

If a source is sufficiently compact, it can become optically thick to its own synchrotron emission due to synchrotron self-absorption. This changes the observed spectral energy distribution from the intrinsic optically thin synchrotron spectrum \(f_s \propto \nu^\alpha\), with \(\alpha \approx -0.7\) as typical value, to \(f_s \propto \nu^{p/2}\), censoring some of the observables by hiding spectral structures and effectively removing access to the low-energy electrons. The self-absorption limited maximum brightness temperature scales as \(T_{\text{max}} \approx 10^{13.5} K (\nu/1 \text{ MHz})^{1/2} (B/1\text{mT})^{1/2}\) (Begelman et al., 1984, eq. A8). Typical magnetic fields in lobes are estimated to be of the order of a few nanoTesla. Lobe surface brightnesses are well below \(10^{12} K\), the maximum allowed by “inverse-Compton cooling catastrophe” (Kellermann and Pauliny-Toth, 1969). The lobes of radio galaxies are typically not affected by optical-depth effects.

This is, however, not true for the “hotspots”, the terminal shock front of powerful jets. These frequently show a spectral turnover at frequencies below 100 MHz. From the size of a hotspot and the turnover frequency one can deduce reasonable limits of the magnetic field strength and power of the jet feeding it.

The inflation of lobes by jets means that mechanical work is done on the surroundings of the radio galaxy. Jets and lobes therefore provide a means for a mechanical coupling between accreting black hole at the centre of the radio galaxy and its surroundings. In current models of galaxy and black hole formation and evolution (e.g. Bower et al., 2006; Croton et al., 2006; Hopkins et al., 2007), this AGN feedback is key to regulating the growth of the black hole and establishing the locally observed correlation between black-hole mass and large-scale galaxy properties such as bulge mass and bulge luminosity (Ferrarese and Merritt, 2000; Gebhardt et al., 2000; Haring and Rix, 2004).

To calibrate and verify these galaxy formation models and to understand the ICM thermodynamics, it is very important to determine the magnitude of this mechanical energy input into a radio galaxy’s surroundings. Since the total energy of the radio lobes is dominated by the lowest-energy electrons, low-frequency observations are essential. Yet, today, still very little is known about the shape of the electron energy distribution at frequencies below 10-100 MHz.

3.2.2. High-redshift galaxies and quasars

One successful method to find high-redshift radio galaxies quasars in the past was to search for ultra-steep-spectrum radio sources (Roettgering et al., 1997). Since observations at lower frequencies include preferentially steeper-spectrum sources, a ULW survey might turn up a large number of high-redshift radio sources. Extrapolating (perhaps optimistically) from source counts at 74 MHz in the VLSS (eq. 21), our example array from the same section (300 elements, 100 km maximum baseline, 10 MHz observing frequency) would thus be able to discover 3 Million sources down to the confusion limit of 65 mJy at 5σ significance over half the sky (500 000 per steradian), which it would be able to survey within roughly two years. For reference, Roettgering et al. (1997) found 30 radio galaxies at \(z > 2\) starting from a sample of \(\approx 700\) ultra-steep spectrum sources.

Another, not yet well-explored method to identify active black holes and jets in the early universe could be to look for young jets that are still contained within the just forming galaxy. They are expected to show up as compact radio sources with synchrotron self-absorption turnover frequencies that are unusually low for their sizes (Falcke et al., 2004). In the local universe these sources are known as Compact-Steep-Spectrum (CSS) and GHz-Peaked-Spectrum (GPS) sources (O’Dea, 1998), with turnover frequencies in the range 100-1000 MHz. At \(z = 10\) these would be shifted to 10-100 MHz and could be readily identified by an ULW array. Interestingly, the only radio-loud quasar so far identified at \(z > 6\) may be a CSS source (Frey et al., 2008).

The number-counts estimate above, of course, includes all kinds of radio sources, so that the 500 000 sources per steradian will also include objects from this category and the ones described next.

3.2.3. Fossil radio galaxies

Radio galaxies and other AGN are now believed to be normal galaxies in a phase of more or less strong accretion onto their central black hole (Croton et al., 2006). It is an important question how long this active phase lasts, and on which timescale their activity recurs.

In the so-called “double-double” radio galaxies with two or even sets of radio lobes (e.g. Schoenmakers et al., 2000), the recurrence is directly visible. From synchrotron ageing arguments, the age of the outer set of lobes is inferred to be of the order of several tens of millions of years, while the inner lobes are inferred to have been launched after an interruption of several to several tens of millions of years. Thus, both the duration of an activity phase and the recurrence timescale are likely within an order of magnitude of \(10^7\) yr.
However, double-double sources are rather rare, perhaps implying that special circumstances might be required for their formation. Other methods are needed to determine the duration of a galaxy’s active phase in the general case. A very direct way of doing so is counting how many “fossil” radio galaxies there are, i.e., those galaxies whose activity phase has switched off recently. However, the fossil galaxies have been elusive so far, in part because the emitted power of a radio galaxy drops as it grows in linear extent even for constant jet power (Blundell et al., 1999). This creates a “youth-redshift degeneracy” in current flux-limited low-frequency radio surveys: only fairly young radio galaxies (ages $\sim 10^9 - 10^7$ yr) are observable at high redshifts ($z \gtrsim 0.8$), while at lower redshifts, radio sources are observable in principle up to ages of $10^9 - 10^7$ yr (Blundell et al., 1999, Fig. 17). Because of this degeneracy no strong statements can be made yet about radio source lifetimes based on the study of low-frequency samples.

Moreover, the only known radio galaxy relics have been identified as cool holes in the X-ray emission of the hot cluster gas (e.g. McNamara et al., 2001; Fabian et al., 2006), not via direct searches for their radio emission (see also the following section on galaxy clusters). Present observations simply may have been done at frequencies that are too high to detect fossil radio galaxies in high-frequency all-sky surveys. This is likely to be the case also with the much better sensitivities of the upgraded E-VLA and E-MERLIN arrays. Lower-frequency observations are better suited for detecting relict radio sources directly because they probe electrons with lower energies, which are those with the longest radiative loss times. This is because of two key properties of synchrotron radiation:

(i) The frequency at which a relativistic electron of energy $E$ in a magnetic field of flux density $B$ emits synchrotron radiation scales as

$$v_c \propto E^2 B,$$

i.e., lower-frequency emission is generated by lower-energy electrons.

(ii) The energy loss of a synchrotron-emitting electron scales as

$$- \frac{dE}{dt} \propto E^2 B^2,$$

i.e., synchrotron cooling times are inversely proportional to energy.

But it is not only the case that synchrotron cooling times increase with decreasing observing frequency – more strongly yet, there is a maximum energy that a synchrotron-cooling electron can have after a given time has elapsed since its acceleration even if it had been given an infinite amount of energy initially (van der Laan and Perola, 1969; Jester et al., 2001), because eq. (29) integrates to $1/E(t) - 1/E(t_0) \propto t - t_0$.

Hence, observations at lower frequencies than currently accessible allow fossil radio galaxies to be detected for a longer interval between the switching-off time and the time of observation, with the accessible radio-lobe lifetime increasing roughly inversely proportional to the observing frequency. Low-frequency radio observations may even be the only possibility for detecting relict radio galaxies. Even the continued non-detection of such fossil galaxies would yield new physical insights, since it would put the strongest constraints on how rapidly radio lobes have to dissipate all of their energy after the central engine ceases to be active.

3.2.4. Galaxy clusters

Just like the gas in radio galaxies, the plasma that fills galaxy clusters emits synchrotron radiation, as well as X-rays through inverse-Compton scattering. The origin of the radiating relativistic particles is unknown. Theories for their origin can be constrained by low-frequency observations, analogous to those of radio galaxies. This will yield important constraints on the formation history of the clusters and of their magnetic fields. Cassano et al. (2006, 2008) have shown that the steep spectra of cluster halos mean that observations at lower frequencies are much more likely to detect the cluster halos, specifically for low-mass and high-redshift objects. Indeed, Brunetti et al. (2008) report the detection of diffuse steep-spectrum emission from the cluster Abell 521 at 240-610 MHz that is not detectable at 1.4 GHz. Again, we stress that even a fairly moderate array of 300 dipoles and 100 km maximum baseline operating at 10 MHz would be able to discover hundreds of thousands of sources in a matter of a few months.

Independently of the question of the impact of AGN feedback on galaxy evolution discussed in §3.2.1 above, the mechanical energy input into the intra-cluster medium (ICM) from giant radio sources is an important quantity in studying the thermodynamics of galaxy clusters, and in elucidating how energy is transported through them in order to establish the observed temperature and entropy profiles. For example, the detailed X-ray observations of the Centaurus and Perseus clusters by Fabian et al. (2005, 2006) show that bubbles blown by the central radio source play an important role. In the case of the Perseus cluster, the X-ray images have revealed sets of “ghost bubbles” without detectable high-frequency radio emission, which are identified as relict sources as discussed in the previous section. These ghost bubbles rise buoyantly in the cluster atmosphere and are therefore part of the energy transport mechanism in clusters. With low-frequency observations, it may be possible to detect radio emission from known bubbles, or new bubbles in clusters that have not or cannot be observed to similar X-ray depth as Centaurus and Perseus. In either case, low-frequency radio observations would yield the most stringent constraints on the ages of such bubbles.

3.3. Galactic/ISM surveys

In this section we discuss how to infer the structure of the interstellar medium and the origin of cosmic rays through low-frequency observations. Beyond these, also the study of pulsars may benefit from these because a confusion limit of 65 mJy at 10 MHz with 1' resolution would be matched to the flux limit of LOFAR for the $\nu^2$ spectrum typical of pulsars. This implies that pulsars could either be detected directly, or limits can be set on their low-frequency cutoff. However, the impact of
temporal broadening in the ISM (§2.4.6) on the low-frequency observability of pulsars needs to be assessed in more detail than we can achieve here.

3.3.1. Structure of the ISM – the Solar System’s neighbourhood

The structure of the ISM in general, and of the supernova-inflated “local bubble” within which the solar system resides, has been inferred by considering absorption lines in ultraviolet and X-ray spectra of stars and the ISM (Frisch, 1998; Breitschwerdt, 2001; Oegerle et al., 2005; Henley et al., 2007, e.g.) and the distribution of pulsar dispersion measures in the sky (Cordes and Lazio, 2002).

Also low-frequency radio observations have been brought to bear on this problem: Peterson and Webber (2002) have modeled the clumpiness of the free electron distribution in the warm ionized medium (WIM) phase of the ISM by matching the expected turnover in the spectrum of the Galactic synchrotron emission to that observed by the RAE-2 (see §1) and IMP-6 (Brown, 1973) satellites. Their method models the radiative transfer through the WIM and requires only knowledge (or reasonable assumptions) about the intrinsic, unabsorbed shape of the underlying synchrotron emission spectrum in addition to the observed brightness distribution. It is applicable over the frequency range where the WIM is neither completely transparent nor completely opaque, i.e., from 0.1–10 MHz according to Peterson and Webber (2002).

This modeling can be extended to infer not only the clumpiness of the ISM, but also the full three-dimensional structure of the WIM. Figure 14 shows the dependence of the ISM “visibility” (distance to the τ = 1 surface) as function of Galactic latitude for different frequencies, and as function of frequency for different wavelengths. The limiting power-law for the line-of-sight distance as function of frequency (right-hand panel in Fig. 14) is given by the inverse of the free-free-optical depth (eq. 11) for the electron density and temperature in the midplane of the Galaxy; for the parameters given in the figure and at wavelengths between 0.1 and 1 MHz, we obtain

\[ D(\tau = 1) \approx 100 \text{ pc} \left( \frac{n_e}{0.132 \text{ cm}^{-3}} \right)^{-2} \left( \frac{T_e}{7000 \text{ K}} \right)^{3/2} \left( \frac{\nu}{1 \text{ MHz}} \right)^2 \] (30)

By scanning through these frequencies and modelling the observed emissivity as function of Galactic coordinates, the structure of the WIM in the solar system’s neighbourhood can be determined in 3D. The requirements on resolution and sensitivity are rather modest: taking 100 pc as a typical ISM bubble size and 1 kpc as a typical distance leads to a degree-scale resolution requirement. The sensitivity requirement is comparatively modest also, since the ISM emission itself sets the background temperature, so that a sensitivity of order of one part in 1000 of the background level is sufficient, i.e., the requirement is \( T_{\text{RMS}} \approx 10^4 \text{ K} \) near 1 MHz.

Around the turnover frequency of 2 MHz, the sky is uniformly foggy; below it, the τ = 1 distance scales as given by eq. (30). Above it, the ISM becomes transparent at lower and lower angles above the midplane as the frequency is increased, i.e., the turnover frequency decreases with the angle above the midplane. Around 2.5 MHz, the entire Galaxy is transparent, while at 0.3 MHz, a frequency well above the putative lunar ionosphere, one would see out only to a few tens of parsec, and see the Galactic plane in absorption against the brighter WIM emission. At these low frequencies, the visibility will therefore be set by the actual substructure of the ISM (supernova shells and bubbles) rather than by the global structure.

On the upside, the short distances out to which we can see at low frequencies imply the effects of ISM angular and temporal broadening are much less severe than for extragalactic observations at higher frequencies. For example, from the NE2001 model (Cordes and Lazio, 2002), the pulse broadening time at 1 MHz over a distance of 500 pc in the direction of the galactic pole is about 90 s, instead of a few days for extragalactic observations at the same frequency. Similarly, the angular scatter broadening is 0.5' instead of ∼30', affording a linear resolution of 0.1 pc at the τ = 1 distance of 500 pc (see Fig. 15 below for an indication of how angular and temporal broadening scale with distance and frequency).

3.3.2. Origin of cosmic rays

Following the overview by Duric (2000), this section describes an approach for elucidating the origin of cosmic rays, i.e., relativistic electrons and protons, with the help of a low-frequency radio array.

The essence of the approach suggested by Duric is to make use of the fact that Galactic HII regions are optically thick at frequencies below about 30 MHz. Therefore, radio emission observed towards such HII regions at low frequencies predominately arises from material along the line of sight to the HII region, i.e., the synchrotron emission from cosmic-ray electrons. By observing the synchrotron emissivity towards different HII regions at different distances from the Earth, a 3D map of the electron density distribution can be built up. To remove the degeneracy in the synchrotron emissivity between electron density and magnetic field strength, the radio data can be compared with gamma-ray observations.

Supernova remnants (SNRs) are known to accelerate high-energy particles through the first-order Fermi shock acceleration mechanism. However, the so-called “injection problem” has not been solved, as the Fermi mechanism can boost particles which are already mildly relativistic to very high energies, but it cannot accelerate thermal particles to relativistic energies. Since the frequency of synchrotron emission scales with the particle energy, low-frequency radio observations can be used to trace the energy distribution of the lowest-energy particles. This is necessary to constrain theories which attempt to solve the injection problem.

The level of turbulence generated by SNRs in the interstellar medium is also relevant to the generation and propagation of cosmic rays. By observing the scatter broadening of extragalactic background sources at various distances from a SNR, the level of turbulence can be probed and mapped in 3D.

However, the resolution required to clearly resolve HII regions and SNRs is of order 1", and the required sensitivity is 0.1 mJy/square arcsecond. Both are a challenge for a lunar
array and may be better done by a ground-based telescope at higher frequencies and resolutions.

3.4. Transient radio sources: suns and planets

Transient sources are particularly interesting in astrophysics because they often indicate very high-energy phenomena, e.g., Gamma-ray bursts (GRBs), supernovae and other stellar explosions, outbursts of accreting black holes, neutrons stars and white dwarfs, reconnection events in the solar corona, or the impact of electrons from the solar wind on planetary atmospheres. Other sources show transient activity simply because of rotation, most notably radio pulsars. Finally, at low frequencies the possibility for detecting coherent emission processes increases due to the macroscopic wavelengths. To allow the detection of transients, high-resolution high-sensitivity observations have to be possible within the timescale of the transient phenomenon.

3.4.1. Solar and planetary radio bursts

Both the Sun and planets produce radio bursts, the Sun through magnetic activity in its corona, and the planets and their moons through interaction of charged particles in the solar wind with the planet’s or moon’s magnetic fields, as is the case for exoplanets (see section 3.4.2). In the case of the Sun, the most violent coronal events give rise to coronal mass ejections (CMEs), disruptions in the solar wind in which large numbers of charged particles (i.e., cosmic rays) are accelerated. The magnetic activity and the charged particles can have dramatic effects on the Earth and Moon, posing a danger to astronauts and electronic equipment outside the Earth’s atmosphere, or disrupting radio communications. These effects have been called “magnetic storms”, and efforts to predict the strength and timing of CMEs impacting on the Earth-Moon system are known as “space weather prediction”. The CMEs themselves are traceable via their low-frequency radio emission as they propagate towards the Earth with an array capable of performing imaging at high time resolution. Since the emission frequency is dependent on density and thus radius, observations down to 100 kHz would trace emission out to about half an AU (using the density model of Mann et al., 1999, e.g.).

While the Sun is the strongest source of radio emission at high frequencies (at 100 MHz and above), bursts from the magnetized planets, in particular Jupiter, can reach a strength comparable to that of the solar bursts or more, i.e., a few MJy (Zarka, 2007). The emission region of Jupiter is very small and not resolved on lunar-like baselines (see Nigl et al., 2007). Hence, an imaging application for Jupiter radio emission is not a scientific driver for a lunar telescope. However, since the emission is strong enough, a single broad-band receiver covering frequencies \( \lesssim 30 \text{ MHz} \) for spectral measurements with high frequency...
and temporal resolution of solar and Jupiter bursts could be envisaged for prototyping and background measurements. For angular discrimination between different burst sources, a small interferometer with comparatively modest baselines (beginning at 500 m for the highest frequency of 30 MHz) would be sufficient. In the long term, also interferometry on a Moon-Earth baseline could be considered to constrain Jupiter bursts in new ways.

3.4.2. Extrasolar planets

While Jupiter’s emission is an additional foreground and interference signal that likely needs to be avoided for high-sensitivity low-frequency observations, it implies that low-frequency radio observations offer the most favourable achievable intensity contrast between an extrasolar planet and its sun, and therefore open up the possibility of a direct detection of extrasolar planets and their magnetosphere.

Zarka (2007) has reviewed the radio emission from extrasolar planets and the results of searches to date. Magnetized planets emit bursts of low-frequency radio waves through the cyclotron maser instability (CMI), in which electrons from the stellar wind interact with the planet’s magnetic field. The CMI is 100% circularly polarized, while solar/stellar bursts are usually unpolarized, offering a mechanism of distinguishing between stellar and planetary radiation. Without a polarization-sensitive array, planetary and solar bursts would be distinguishable by temporal structures, either within bursts or from modulations on the timescale of the planetary orbit. Lazio et al. (2004) give 15 minutes as typical burst duration, while Zarka (2007) shows an example of a burst with an overall duration of several hours, with frequency structure down to the millisecond level (in his nomenclature, only the features on sub-second timescales are called “bursts”).

Since the CMI is only active in the presence of strong magnetic fields, either in the planet or the stellar wind, this mechanism is most suitable for the detection of magnetized Jupiter-like planets. Earth-like planets will only give rise to radio bursts if they are embedded in a stellar wind which is itself strongly magnetized. Even if this is a rare situation, the radio emission might be the most readily detectable signal of such planets.

A crucial parameter is of course the rate at which planetary bursts occur, and hence the survey time that needs to be invested in order to obtain a statistically relevant sample of burst observations. However, very little is known about the likely recurrence rate, which depends on the unknown way in which CMI bursts are triggered. One possibility is that they have “light house beams” similar to pulsars, with intrinsically continuous emission that is confined to a solid angle and bursts being observed when the emission beam sweeps over the telescope (e.g., Zarka et al., 2004).

Another possibility is that CMIs are triggered or enhanced by coronal mass ejections (CMEs; see Gurnett et al., 2002, who give an example of a solar CME enhancing Jovian radio emission). In this case, the relevant timescale is the CME recurrence timescale modified by the probability that a CME in fact interacts with the planet. Based on such considerations, Khodachenko et al. (2007) expect CME rates in excess of the solar value of about 6 CMEs/day, in particular for M dwarfs which are intrinsically more active than solar-type stars. Grießmeier et al. (2007a) add that CMEs can lead to an effectively continuous enhancement of planetary radio emission, in particular in young stars with their frequent CMEs.

In spite of the current uncertainties about the rate of exoplanetary low-frequency bursts, the possibility of a direct detection of exoplanets via such bursts is an exciting prospect. Being closer to their star than Jupiter is to the Sun, hot Jupiters should show much more intense radio emission than Jupiter itself, at least of order 10^5 times more (Lazio et al., 2004; Grießmeier et al., 2005; Zarka, 2007). In particular, Lazio et al. (2004) have computed the expected radio emission for the extrasolar planets known at the time (mostly hot Jupiters) and found that they should have bursts of flux densities close to 1 mJy at frequencies of several tens of MHz up to a few GHz, while a few lower-mass planets should emit at frequencies around 0.1-1 MHz and reach tens to hundreds of mJy. Recently, Grießmeier et al. (2007b) have compared predictions for three separate emission processes for currently known exoplanets, and find lower peak frequencies and average emitted powers than Lazio et al. (2004).

Referring to Fig. 4, the detection of a 1 mJy burst at 5σ significance over a typical duration of 15 minutes (Lazio et al., 2004) at a frequency of 30 MHz and at 25% bandwidth requires of order 10^5 dipoles, while detecting a 10-mJy burst of equal duration at 3 MHz (and again with 25% bandwidth) still requires about 10^6 dipoles; at fixed frequency, the required number of dipoles is inversely proportional to the required sensitivity. Source variability and their Galactic nature would help to overcome the confusion limit. The requirement for the number of dipoles is driven by the need to detect a burst within its duration. Furthermore, there is no need to map the spatial structure of a burst, i.e. the sensitivity requirement is on flux, not surface brightness. Hence, the dipoles can be distributed over baselines that yield arcminute-scale resolution, which is sufficient for establishing positional correspondence to observations at other wavelengths.

Before investing in the construction of such an array, it is important to verify that the temporal pulse broadening due to ISM turbulence (§2.4.6) does not smear out an exoplanetary burst with a duration of ≈15 minutes beyond recognition. We used the Fortran implementation of the NE2001 model for the distribution of free electrons in the Galaxy by Cordes and Lazio (2002) to determine the scattering regime, pulse-broadening times, and angular broadening for 176 exoplanets.11 The planets have a mean distance of 45 pc, and sample a large range of Galactic longitudes and latitudes (though their distributions naturally follows that of Galactic stars, so that they are concentrated towards the Galactic plane).

Figure 15 shows the pulse-broadening times and angular broadening scales for the exoplanets as a function of distance. Although all exoplanets are in the strong scattering regime at frequencies below 30 MHz, the angular broadening even at

11 Taken from the list at http://exoplanets.org
1 MHz reaches at most a few 10s of arcseconds, which is negligible compared to the likely resolution of a lunar array (compare Fig. 5). The pulse broadening times at 1 MHz are nearly exclusively well below 1 ms, and in all cases less than 10 ms. At 1 MHz, the pulse broadening is a factor of $10^{1.4}$ larger (eq. 16), but still below 1 sec and hence negligible for all planets up to 70 pc distance. Fewer than 1/6 of the known planets are at a greater distance than 70 pc, and only 5 of the 113 planets considered by Lazio et al. (2004) are predicted to emit at frequencies below 1 MHz because they considered only Earth-based radio observatories. Moreover, as noted above, there are indications (P. Zarka, priv. comm.) that the pulse broadening may increase less rapidly with wavelength than expected from eq. (16). But even if that equation holds, Fig. 15 shows that angular and pulse broadening by the ISM impose no significant constraints on the observability of exoplanetary radio bursts.

3.5. Ultra-High Energy particle detection

3.5.1. Hadronic Cosmic Rays

Ultra-high energy cosmic rays (UHECRs) are a rare population of elementary particles, likely composed of protons, atomic nuclei, and potentially photons or neutrinos (Nagano and Watson, 2000). Their energies can reach much higher values than those currently achievable in man-made particle accelerators. The total flux of cosmic rays is about 1 particle m$^{-2}$ sterad$^{-1}$ year$^{-1}$ above a fiducial particle energy of $10^{16}$ eV. The cumulative flux spectrum$^{12}$ drops roughly as $E^{-2}$ with higher energy thresholds, flattening to $E^{-1.75}$ below the so called “knee” in the cosmic ray spectrum at $2 \times 10^{15}$ eV (Hörandel, 2003). Above the knee, the cosmic ray flux is roughly given by

$$F_{CR}(E > E_c) = 30 \text{ particle km}^{-2} \text{ sterad}^{-1} \text{ day}^{-1} \left(\frac{E_c}{10^{15}\text{eV}}\right)^{-2}$$

at $E_c > 2 \times 10^{15}$ eV.

At the highest energies, the particle spectrum extends up to and perhaps beyond $10^{20}$ eV. The currently best information on the spectrum is available from the Pierre Auger Observatory (The Pierre Auger Collaboration, 2008), showing the well-known “ankle” around $4 \times 10^{18}$ eV and a steepening beyond $4 \times 10^{19}$ eV. The latter is likely due to the Greisen-Zatsepin-Kuzmin (GZK) cutoff, a suppression of extragalactic UHECRs due to proton-gamma interactions with the cosmic microwave background. Origin and nature of UHECRs is still largely unknown, but a link with nearby accreting supermassive black holes has been made (The Pierre Auger Collaboration, 2007).

A major obstacle to understanding these most energetic particles in the universe is their small flux of less than 1 particle per

---

$^{12}$i.e., the total number of particles above that energy
UHECRs can be detected through their intense particle cascades which they initiate when encountering a target. For example, if UHECRs hit the lunar surface a particle shower will pass through the lunar regolith with a speed greatly exceeding the local speed of light (the refractive index of the lunar regolith is $n = 1.73$, see Alvarez-Muñiz et al. 2006 and references therein), leading to intense Cherenkov light (Zas et al., 1992; Dagkesamanskii and Zheleznykh, 1989; Hankins et al., 1996; Gorham et al., 2004). Due to the finite length of the shower of a few meters, the emission must be coherent at long radio wavelengths. The shower will then show up as a coherent ultrashort radio burst on a time scale of $\Delta t = 1/\Delta v$, which is tens of nanoseconds up to microseconds in the low-frequency regime ($v \approx 1 \sim 100$ MHz).

Hence, cosmic rays will in principle create an impulsive noise background on top of the steady Galactic noise, but the emission will in turn also allow one to them in more detail. Falcke and Gorham (2003) and Scholten et al. (2006) have shown that the radio emission from such lunar showers is detectable at the very highest energies by a LOFAR-like telescope even from Earth. Especially for cosmic rays and neutrinos at energies $> 10^{21}$ eV, well beyond the GZK cut-off, Earth-based low-frequency observations of lunar radio bursts generated by interactions within the moon offer a unique possibility to study these elusive particles.

However, interpretation of these results in the end depends on subsurface properties of the moon and shower-development at these high energies (e.g., due to the Landau-Pomeranchuk-Migdal effect), making in-situ measurements invaluable. This is impossible on the Earth since accelerators cannot produce particles at such high energies and the Earth’s atmosphere prevents UHECRs from directly hitting the surface.

Moreover, placing low-frequency antennas on the moon would have several interesting advantages compared to Earth-bound radio experiments: First, being 400,000 times closer to the location of the shower (assuming a fiducial shower distance of 1 km on the moon) than the Earth, would, roughly speaking, increase the sensitivity to cosmic rays and neutrinos by a factor 400,000$^2$ and allow correspondingly lower-energy particles to be detected. Secondly, lunar radio pulses observed from the ground get dispersed by free electrons in the Earth ionosphere which can be avoided with in-situ measurements. Also, the expected radio-quietness at the lunar surface would greatly aid in solidly detecting these events.

Using the equations given by Scholten et al. (2006), we find that a single dipole antenna operating at 100 MHz with 20 MHz bandwidth should be able to detect of order 7 cosmic-ray events per day at $10^{18}$ eV and above at flux densities of 100 MJy, corresponding to signal-to-noise levels of 10 or more, assuming events at a distance of or less 1 km, arriving from a cone of half-opening angle 45° (Fig. 16).

One advantage of placing the radio receivers directly on the moon instead of on the Earth is that high-energy events, which are rarer, can be detected over greater distances from the location of the antennas. Therefore, a larger fraction of high-energy
events is observed, and the differential energy distribution of the observed events scales as $1/E$, even though the intrinsic distribution scales as $1/E^2$; the difference arises because more energetic events produce quadratically greater fluxes, which are detectable over linearly greater distances, increasing the collecting area quadratically with energy. Thus, a lunar array finds a larger fraction of the rarer high-energy events than an Earth-based array.

The total rate of detected events increases logarithmically with energy, up to a maximum given by the largest distance out to which an event’s radio pulse is still visible to the antenna(s). The total event count will still be heavily dominated by low-energy events, so that the total number of detected events above some energy is approximately given by the number of events at that energy which occur within the radius for which the signal-to-noise ratio $S/N$ meets the detection threshold.

At least three antennas are necessary to determine the location of a cosmic-ray event by triangulation and hence the original energy of the particle. An array of three elements will achieve an $S/N$ of about 2.5 times that in Fig. 16, since the $S/N$ scales as $\sqrt{N(N-1)}$, where $N$ is the number of antennas. The total event rate above some fixed $S/N$ scales as $\ln(N(N-1))$, because lower minimum energies $E_{\text{min}} \propto 1/\sqrt{N(N-1)}$ become accessible and the differential event rate scales as $1/E_{\text{min}}$.

If the observing frequency is lowered from 100 MHz to 10 MHz, the flux density per event will be much lower, while the sky background is slightly higher, and, of course, the available bandwidth is smaller. All three effects mean that the $S/N$ for a fixed event energy is smaller at lower observing frequencies. On the other hand, the emission is expected to be more isotropic at lower frequencies; while only events above $10^{20} \text{ eV}$ would be observable at 10 MHz, all such events are detectable out to a radius of 5 km by a single dipole, with a rate of one detectable event about every 2 years. This means that observing cosmic rays is feasible in the ULW region, and an astronomical array of $\approx 100$ elements with 5 km spacing could detect of order 30 events/year at $E \gtrsim 10^{20} \text{ eV}$. This is in principle competitive with the largest cosmic ray detectors today.

3.5.2. Neutrinos

The arguments used for deriving the radio flux generated by hadronic cosmic-ray events also apply to neutrino interactions, so that the basic array parameters are identical for the detection of both neutrinos and more massive cosmic rays. The basic difference to the hadronic component of UHECRs is that neutrinos will penetrate deeper into the moon before they interact. The mean free path is $130 \text{ km} \left( E_{\nu_e}/10^{20} \text{ eV}\right)^{1/3}$ and hence the radio pulse would come from deep below the surface. Here, lower radio frequencies have the additional advantage of penetrating deeper into the moon crust, thereby enlarging the usable detector volume. Scholten et al. (2006) quote an attenuation length of

$$l = 9 \text{ m} \left( \nu/\text{GHz}\right)^{-1}. \quad (32)$$

If this scaling continues all the way to 10 MHz, then a single dipole can see a volume of $\approx 1.5 \text{ km}^3$ of detector material for neutrino interactions. Hence, a simple radio array could in principle easily cover several tens of cubic kilometers. This would be more difficult at these wavelengths on Earth, due to the global short-wave radio interference. However, neutrino searches have already been proposed and performed over Antarctica at higher frequencies (Barwick et al., 2006; Halzen, 2007; Spiering, 2007). Therefore, the competitiveness of a lunar array needs to be evaluated more thoroughly. It may also be that the difference in geometry of the moon and the density of the lunar crust leaves an interesting parameter space open.

These experiments could run in parallel with an imaging array. They would require an additional signal path with a pulse detection algorithm and a raw time series buffer of several microseconds to milliseconds per antennas. The trigger event time has to be at the tens of nanosecond level. Unlike for most other applications, the antennas would require a large acceptance in the horizontal planes to detect “grazing-incidence” pulses and toward the ground. For simple dipoles without ground plane an up-down ambiguity remains; however, bandwidth-limited pulses produced in the near field above the lunar surface appear exceedingly unlikely, so that this up-down ambiguity should not be a concern. The distinction between neutrinos and protons/nuclei in the cosmic radiation would be made based on the geometry: only neutrinos would come from well below the array.

As a side-effect of a lunar neutrino detector, one could observe the Earth as a neutrino source. The emission is due to the production of atmospheric neutrinos which are generated in particle showers initiated by cosmic rays hitting the Earth’s atmosphere\(^{13}\). Since the terrestrial neutrino flux is well known, one could wonder about using the Earth-Moon baseline for neutrino oscillation measurements.

3.6. Meteoritic Impacts

Finally, we point out the possibility that impulsive radio noise measurements might also be a potential tool to investigate meteoritic impacts on the moon. Such impacts seem to produce optical flashes which are visible even from Earth (Ortiz et al., 2000; Cudnik et al., 2003; Cudnik, 2007). It is not inconceivable to think that such impacts are also accompanied by low-frequency radio emission. This may be in analogy to strong electromagnetic pulses (EMPs), known from nuclear explosions, which have already been used to explain radio emission from cosmic rays in the Earth’s atmosphere (Colgate, 1967). One could also imagine an interaction between the explosion and local electrostatic or magnetic field, however, all of this remains speculative at the moment.

At least, radio emission from impacts has already been used by radio experiments in free space, where radio antennas ex-

\(^{13}\)It is amusing to note that the level of neutrino emission from a rocky planet like Earth is in fact a signal for the presence of an atmosphere. Showers in a solid medium that is not shielded by an atmosphere produce far fewer neutrinos due to the much shorter shower length. Unfortunately the level of atmospheric neutrino emission is far too low to be used in the study of exoplanets.
hibit radio pulses due to high-speed impacts of dust grains. For example, based on early findings of Voyager II, a radio antenna on the Cassini spacecraft was used to measure dust particles from Saturn’s rings hitting the entire spacecraft (Kurth et al., 2006). The idea here is that the energy of the impact creates an expanding ionized plasma with an ambipolar electric field, which in turns creates a strong voltage pulse on the antenna or spacecraft. The time scale of the pulse is on the ms scale and hence should be well discernible from CR impacts. Moreover, the frequency spectrum of the radio pulse seems related to the size of the impacting particle. The radio signals might therefore reveal the size and mass distribution of the particles. However, this effect typically is associated with electrostatic noise generated at or near the antenna itself. In how far this effect can applied to a lunar antenna remains unclear at present.

4. Summary & Discussion

In this paper we have summarized some of the main science goals, requirements, and limitations for astronomy at very low radio astronomy frequencies below 100 MHz, i.e., ultra-long wavelengths (ULW). In this frequency range the moon, in particular its far side, offers by far the best (albeit not cheapest) wavelengths (ULW). In this frequency range the moon, in particular its far side, offers by far the best (albeit not cheapest) site for conducting observations. In several cases a lunar telescope even offers unique insights.

The main limitation for ground-based low-frequency telescopes is degradation of the image quality due to the Earth’s ionosphere in the 10-100 MHz frequency range. At very low frequencies (< 10–30 MHz), the ionosphere is even blocking our view of the universe completely. Space and in particular lunar radio arrays are therefore back in the focus. Here we have discussed a number of science areas where this frequency range may be of special interest: cosmology with primordial HI at z = 10–100, extragalactic surveys of radio galaxies, clusters and galaxies, astroparticle physics of cosmic rays and neutrinos hitting the moon, transient flares from stars and exoplanets, the 3D structure of the local interstellar medium, and solar system objects.

4.1. Science goals

Table 1 summarizes the various science goals and requirements. Clearly of greatest interest is the question of the dark ages and HI observations at the highest redshifts. This would contain a huge amount of original information about the very early universe unaltered by the complicated astrophysical processes associated with cosmic reionization.

For example, a determination of the exact boundaries of the Epoch of Reionization through its global signal could in principle be achieved with a single antenna in a dark crater near one of the lunar poles. This would provide almost perfect shielding from any radio interference and temperature (hence gain) variations – ideal conditions for such measurements. However, this experiment is also being tried from the ground under somewhat less than ideal conditions and the outcome has to be awaited (e.g., see the loose upper limit just obtained by Bowden et al. 2008 with the EDGES experiment).

The next step up is the measurement of the HI power spectrum above z ≳ 10 and into the dark ages. Since the HI gas will trace directly the dark matter distribution, this is an excellent way to measure the primordial density fluctuations which led to the formation of large scale structures in the universe.

Measuring the baryonic acoustic oscillations (BAOs) also during the dark ages has several advantages, the most important one being the lack of non-linear processes and biases. Non-linear effects, e.g., the rapid growth of ionization bubbles from the first stars, affect HI measurements during the epoch of reionization, while unknown star formation efficiencies affect measurements at lower redshifts using galaxies. Hence, measurements of BAOs in the dark ages could tackle, for example, the issue of non-Gaussianity in the density fluctuations (Cooray et al., 2008). This is a quantity directly revealing information about the inflationary phase of the universe. Measuring the BAOs at z = 50 on a scale of tens of arcminutes requires of order 10⁵ antennas, because the amplitude of BAOs is at the scale of 10⁻⁸ of microKelvin. If one tries to push the resolution further down to arcminute scales in order to discriminate between cosmological models as discussed by Loeb and Zaldarriaga (2004), one again requires at least of order 10⁵ antennas at z ≥ 20. This is a somewhat futuristic experiment for a space array. If this is realized with 5 m crossed dipoles one requires about 1000 km of wires, which corresponds to of order 1 ton of metal for a light weight wire of one gram per meter. While this is not the limiting factor, the power consumption and the electronics might be. Such broadband low-frequency receivers currently require a power of the order of several Watts, bringing the total energy consumption rate up to several 100 kW plus a correlator – hence this would require some serious but not completely outrageous lunar infrastructure to be available.

The main obstacle for achieving the same science goal from the ground is the ionospheric seeing – not because of the resolution, but because of calibration problems. The detection of the HI signal against the enormous sky background requires a perfect removal of all strong confusing sources and their side lobes. Since low-frequency observations are all-sky observations, the confusing sources will be seen through all the different patches of the ionosphere on the sky, making this a calibration problem with many degrees of freedom. Once more, the coming years have to show down to which frequencies ground-based arrays can calibrate out the obstacles imposed by ionosphere, interference, and foreground. Current experiments do not even attempt to push into this high-redshift range, since well-calibrated observations at frequencies v < 30 MHz corresponding to 21-cm emission from z ≥ 50 seem completely out of the question for ground-based observations.

As the ultimate goal one could imagine an actual imaging experiment of the dark ages, delivering “all one needs to know” about the Universe. However, as we have seen here, the required collecting area is several tens of square kilometers to achieve the necessary surface brightness sensitivity with arcminute-scale resolution. This puts such a project in a very distant future.
## Table 1. Overview of science cases and requirements.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Requirements</th>
<th>Frequency Resolution</th>
<th>Baselines</th>
<th>Expected signal $N$(Antennas)</th>
<th>$t_{\text{exp}}$ (5σ)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cosmology</strong></td>
<td></td>
<td>MHz</td>
<td>Baselines</td>
<td>Expected signal</td>
<td>$t_{\text{exp}}$ (5σ)</td>
</tr>
<tr>
<td>Global Epoch of Reionization$^b$</td>
<td>3.1</td>
<td>50–150</td>
<td>$2\pi$ sr</td>
<td>0</td>
<td>5–50 mK</td>
</tr>
<tr>
<td>Global dark-ages signal</td>
<td>3.1.2</td>
<td>30–45</td>
<td>$2\pi$ sr</td>
<td>0</td>
<td>0–40 mK</td>
</tr>
<tr>
<td>EoR tomography</td>
<td>3.1.3</td>
<td>50–150</td>
<td>$\approx 10'-1'$</td>
<td>0.7–20</td>
<td>$\approx 1$ mK</td>
</tr>
<tr>
<td>EoR 21-cm power spectrum</td>
<td>3.1.4</td>
<td>50–150</td>
<td>$\approx 2^\circ-2'$</td>
<td>0.05–10</td>
<td>$\approx 0.3$–5 mK</td>
</tr>
<tr>
<td>Dark-ages 21-cm power spectrum</td>
<td>3.1.4</td>
<td>30–45</td>
<td>$\approx 20'-2'$</td>
<td>1–20</td>
<td>$\approx 0.03$–1 mK</td>
</tr>
<tr>
<td><strong>Extragalactic surveys</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Galactic surveys</td>
<td>3.2</td>
<td>10</td>
<td>1'</td>
<td>0.1–100</td>
<td>$\approx 65$ mJy</td>
</tr>
<tr>
<td>Extra-galactic surveys</td>
<td>3.3</td>
<td>0.1–10</td>
<td>degrees</td>
<td>0.3–30</td>
<td>10000 K</td>
</tr>
<tr>
<td>Origin of cosmic rays</td>
<td>3.3.2</td>
<td>0.1–30</td>
<td>1'' (3–30)$\times 10^3$</td>
<td>155,000 K</td>
<td>100,000</td>
</tr>
<tr>
<td><strong>Transients</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solar and planetary bursts</td>
<td>3.4</td>
<td>0.1–30</td>
<td>degrees</td>
<td>0.5–200</td>
<td>M Jy</td>
</tr>
<tr>
<td>Extrasolar planets</td>
<td>3.4.2</td>
<td>0.5–30</td>
<td>$\lesssim 1'$</td>
<td>$\gtrsim 35$–1000</td>
<td>1–10 m Jy</td>
</tr>
<tr>
<td><strong>Ultra-High Energy particles</strong></td>
<td>3.5</td>
<td>10–100</td>
<td>N/A$^g$</td>
<td>0–5</td>
<td>100 M Jy</td>
</tr>
<tr>
<td><strong>Meteoritic Impacts</strong></td>
<td>3.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Surface brightnesses given in Kelvin, fluxes in Jansky

$^b$ Here we use $10 < z < 20$ as fiducial redshift limits for the Epoch of Reionization (EoR), and $30 < z < 50$ for the Dark Ages.

$^c$ This requirement is driven by the demanding the necessary surface brightness sensitivity over timescales compatible with the human lifetime, i.e., of order of years. Note the upper limit $N$(Antennas) $< N_{\text{max}} = 4.7 \times 10^7 (\theta/1')^2$ from the shadowing limit (eq. 9). The array parameters for 21-cm tomography and power-spectrum observations scale as $N_{\text{ant}} \sqrt{t_{\text{exp}}} S^{-1} \theta^2 (1+z)^3 = \text{const}$, where $S$ is the desired sensitivity.

$^d$ Confusion limit from eq. (22); about $3 \times 10^6$ sources are expected to this limit in a half-sky survey.

$^e$ Moon diameter: 3.476 km

$^f$ The number of antennas is driven by requiring a 5-σ burst detection within the burst typical burst duration of 15 minutes, with $t_{\text{exp}} N_{\text{ant}} = \text{const}$.

$^g$ Near-field triangulation of event sites helps, but is not strictly required.

$^h$ Expected event rate for hadronic cosmic rays: $\sim 7$/day at $E > 10^{18}$ eV for 1 dipole at 100 MHz; $\sim 30$/year at $E > 10^{20}$ eV for 100 dipoles with 5 km spacing at 10 MHz.
Apart from the fundamental cosmology science goal, other topics are also of potential interest. Clearly, Galactic and extragalactic surveys are needed to shed light on how the Universe looks at these wavelengths. One expects to be particularly sensitive to fossil radio sources in clusters, and to be able to constrain the total energy output of quasars. As shown by the survey equation (24), an extragalactic half-sky survey with arcminute resolution at 10 MHz is possible within two years with an array of a few hundred elements and baselines up to 100 km down to a 5 σ detection of the confusion background.

A fairly unique science application at low frequencies would be the 3D tomography of the local warm interstellar medium. Since the hot gas in our galaxies starts to become optically thick between 0.1 and 3 MHz, a ULW radio array could see the entire spatial distribution of this gas from the Galactic Center to the Local Bubble by simply stepping through all frequencies. The gas will have imprinted in its structure the recent history of the Galaxy and in particular of the solar system environment. This is of particular interest since the Local Bubble pressure has a direct effect on the size of the heliosphere and thus on the cosmic ray flux on Earth and its related effects on weather and biological mutation rates (Scherer et al., 2006). Hence, one might be able to shed some light on the co-evolution of life in the solar system and its interstellar environment. To perform this survey, a lunar telescope would require several tens to hundred dipoles, several tens of kilometer baselines and a broad frequency coverage down to 0.1 MHz. That could be a realistic science goal for the first lunar telescope installations.

Another new application in ULW astronomy is the search for transient sources. Here very little is known a priori. Jupiter is the prime example with its cyclotron maser flares at 10 MHz which are completely invisible at frequencies above 40 MHz due to a sharp spectral cutoff, depressing the flux by 5 orders of magnitude (Zarka, 2007). It has been suggested that this could be used for the detection of the magnetospheres of exoplanets (Lazio et al., 2004; Grießmeier et al., 2005, 2007b), but will certainly also be of interest to investigate solar flares in nearby stars. Transient detection requires good instantaneous (u, v) coverage, tens of thousands of dipoles, some spatial resolution (baselines of 10-100 km), triggering mechanisms, and buffering electronics.

Due to its huge mass and lack of atmosphere, the moon is also a unique ultra-high energy cosmic ray detector. This is currently being used by ground-based experiments which exploit the fact that particle cascades in the lunar regolith should produce strong radio Cherenkov emission. At the very highest energies for cosmic rays this promises by far the most effective detection mechanism (Scholten et al., 2006). The main limit at present are systematic uncertainties of lunar properties. These could be overcome with the help of a relatively small lunar surface radio array which would make in-situ measurements of these cosmic-ray-induced radio flashes. Significant advances might already be possible with a first prototype array of only a few dipoles. Of particular interest is the transparency of the lunar regolith below 10 MHz. If this continues to increase with decreasing frequency, already a few dipoles could provide an effective neutrino detector volume of many cubic kilometers in the future. A by-product of such impulsive radio noise measurements could be the measurement and localization of meteoritic impacts on the moon.

4.2. Open questions

It is obvious that a full-blown lunar ULW radio array with thousands of elements cannot be the first step. Indeed, there are a number of open questions about observability and design constraints for a moon-based ULW radio telescope, which need to be answered before a design for a lunar array can be contemplated in detail, let alone be finalized:

(i) What are the properties of the lunar “ionosphere”, in particular at the likely sites for a lunar array on the far side and near the poles?
(ii) What are the subsurface and electrical properties of the moon?
(iii) How strong is the shielding of terrestrial radio waves in the polar region?
(iv) How strong is the impulsive radio background noise on the surface?
(v) How strong are electromagnetic effects associated with meteorite impacts that could cause impulsive radio noise on the lunar surface?
(vi) Are the IPM and ISM scattering extrapolations from higher frequencies still accurate near 3 MHz?
(vii) What is the detailed synchrotron and free-free emissivity and optical depth of the ISM, as a function of galactic coordinates?
(viii) What is the lowest frequency at which ionospheric fluctuations above a terrestrial observatory can be calibrated sufficiently well to achieve the dynamic range necessary for imaging faint sources?

Some of these questions will be answered by LOFAR or MWA in the near future, followed later by the SKA. Instruments on current-generation lunar orbiters like the on-going Japanese KAGUYA/SELENE mission will also provide new insights into questions about the electrical properties of the moon. Nevertheless, some questions will remain that will require measurements and prototyping on the moon directly.

4.3. Stepping stones en route to a large lunar array

The big advantage of the phased dipole array technology is that it is robust, very flexible, and scalable with the number of antennas. Therefore, one can easily define a program with several steps signified by a growing number of antennas. Dipole antennas can in principle be packaged in a small device and therefore could initially piggy-back on other missions for feasibility studies. The following steps could be envisaged:

(i) a dedicated pathfinder experiment with a single antenna to
   – analyze the transmission of the lunar ionosphere
   – investigate the surface and subsurface electrical properties of the Moon
- determine the beam shape and stability of a dipole on the surface
- monitor the radio noise environment on the lunar surface
- measure the redshift of the global Epoch of Reionization (§3.1.1)

(ii) a two-element interferometer with a moving antenna that can
- demonstrate lunar interferometry
- perform a first simple ULW all-sky survey

(iii) a small lunar ULW telescope consisting of \( \geq 3 \) dipoles in the 100 MHz regime that can
- act as cosmic-ray, neutrino, and meteoritic impact detector (§§3.5.1, 3.6)
- demonstrate lunar interferometry

(iv) an imaging telescope operating in the 10 MHz regime with 30–300 dipoles distributed over 30–100 km, for
- extragalactic surveys of large-scale radio structures in clusters, radio galaxies, etc. (§3.2)
- tomographic mapping of the Galactic ISM (§3.3.1)
- monitoring planetary radio bursts (§3.4.1)
- neutrino and cosmic ray measurements (§3.5)

(v) a full-size array optimized for the 10–100 MHz regime, with \( 10^3 - 10^7 \) antennas spread over tens of kilometers to
- investigate the cosmic dark ages (§3.1)
- search for transient Galactic sources, in particular exoplanets (§3.4.1)

Among the early prototypes, a simple two-element interferometer near one of the poles is a particularly appealing concept. It could consist of a fixed antenna and one installed on a rover, with the rover operating over several months and slowly moving further away. The combination of lunar rotation and motion of the rover out to tens of kilometers would provide a decent \((u,v)\) coverage and allow one to perform an interesting imaging survey already with a small mission.

4.4. Potential realizations

The realization of these arrays is a matter of current study. Dipoles are the preferred primary element for the simplicity and ease of deployment, but also magnetic loop antennas and tripoles are possible options. Since the Moon’s surface is a much better insulator than the Earth’s, in principle the dipoles could be placed directly on the lunar surface, without a defined ground plane (Takahashi, 2003).

To keep data rates for earth-moon telemetry manageable (details below), an array with a large number of elements requires a correlator to be placed on the moon. For the deployment of a large number of antennas, Jones et al. (2007) have considered printing of antennas onto very thin foil that can be rolled out by astronauts. However, for a smaller number of dipoles robotic deployment by rovers or even simple ballistic ejection mechanism seem to be sufficient.

A major concern is the power supply to individual antenna elements. This has to be achieved either locally, through small solar cells and batteries for some fractional night operation, or through wire connections to a central solar power generator. The limited capacity of current rechargeable batteries is a major limitation for any night operation. Day-time operation has always the danger of being affected by strong interference from the sun and the lunar ionosphere at the very longest wavelengths.

A second challenge is the data connection between the individual elements and the central data processor (“correlator”) on the Moon, where the data streams from all antennas have to be correlated with each other, and the connection between the correlator and the ground station on Earth. For 10 MHz bandwidth and 8-bit digitization at 20 MS/s a connection of \( \approx 200 \) Mbit/sec is needed. Rather than digitizing the radio waves at the dipoles and risking self-generated interference, it may be preferable to carry out the digitization close to the correlator. For this, the analog radio signal could be modulated onto a laser or a glass fibre connection. Due to the curvature of the moon and its rocky nature, laser connections will only work over a limited distance. Wireless radio communication at high frequencies could also be considered, but this option is disfavored as it would disturb the radio-quiet nature of the site. The data rate demands between the correlator and Earth are much less than the intra-lunar connections, but have to go through a relay satellite for a location on the far side.

The data correlator itself does not appear to be a fundamental problem. Modern general-purpose Field Programmable Gate Arrays (FPGA)-based correlators (e.g., von Herzen, 1998; Woody et al., 2004; Takeuchi et al., 2005; Parsons et al., 2006, 2007) are already able to integrate the cross-correlation of 30 antennas with 10 MHz bandwidth essentially on a single chip.

4.5. Conclusions

The low frequency end of the electromagnetic spectrum is currently not well-explored, and the very-low frequency end is indeed one of the last portions of the electromagnetic spectrum to remain terra incognita in astrophysics. A lunar low-frequency array would therefore enable pioneering observations in a number of scientific areas (Tab. 1), from the Dark Ages in cosmology to exoplanets and astroparticle physics. The field is still in its infancy and serendipity may hold some surprises. After all, the transition to long wavelengths also marks a qualitative transition, where coherent processes play a much stronger role and hitherto unknown ultra-steep spectrum source could reveal themselves for the very first time. If lunar exploration becomes a reality again in the next decades, then a low-frequency array for astrophysics and space science should be an important part of it, supplementing the impressive array of new radio telescopes on the ground.

Acknowledgements Early parts of this work were supported by ASTRON and EADS Astrium within a joint concept study for a lunar infrastructure for exploration. We are thankful to G. Woan, J. Noordam, A. Gunst, S. Wijnholds, J. Bregmann, M.
Meijers, H. Butcher, M. Garret, D. Werthimer, J. Lazio, H. J. Heidmann, H. Guenther, P. Zarka, S. Zaroubi, and Hartmut Müller, as well as the participants of the lunar observatories workshops in Bremen for many interesting discussions on this and related topics. We thank Rick Perley for discussions on the survey equation and wide-field imaging. Tom Maccarone for the pulsar idea, and Benedetta Ciardi for discussions of the global 21-cm signal from the dark ages and providing the input data for Figs. 10 and 9. We are grateful to the anonymous referee for constructive comments which have allowed us to improve this review.

References


von Herzen, B., Jul. 1998. Digital cross-correlation at 250 MHz

van der Laan, H., Perola, G. C., Dec. 1969. Aspects of radio...

The Pierre Auger Collaboration, Nov. 2007. Correlation of the...

Takahashi, Y. D., 2003. New astronomy from the moon: a...

Visbal, E., Loeb, A., Wyithe, S., 2008. Cosmic ray


van der Laan, H., Perola, G. C., Dec. 1969. Aspects of radio...