Towards nuclear magnetic resonance $\mu$-spectroscopy and $\mu$-imaging

P. J. M. van Bentum, J. W. G. Janssen and A. P. M. Kentgens*

Department of Physical Chemistry, NSRIM Center, University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands. E-mail: Arno.Kentgens@nmr.kun.nl; Fax: +31-24-3652112; Tel: +31-24-3652078

Received 24th March 2004, Accepted 14th June 2004

First published as an Advance Article on the web 5th August 2004

The first successful experiments demonstrating Nuclear Magnetic Resonance (NMR) were a spin-off from the development of electromagnetic technology and its introduction into civilian life in the late forties. It was soon discovered that NMR spectra held chemically relevant information making it useful as an analytical tool. By introducing a new way of detection, moving away from continuous wave spectroscopy, Fourier Transform NMR helped to overcome sensitivity problems and subsequently opened the way for multi-dimensional spectroscopy. As a result NMR has developed into one of the most powerful analysis techniques with widespread applications. Still sensitivity is a limiting factor in the applicability of NMR. Therefore we witness a renaissance of technique development in magnetic resonance striving to improve its receptiveness. This tutorial review introduces the efforts currently made in miniaturizing inductive detection by designing optimal radio-frequency microcoils. A second approach is to introduce a new way of detecting magnetic resonance signals by means of very sensitive micromechanical force detectors. This shows that the detection limits in terms of absolute sensitivity or imaging resolution are still open to significant improvements.

Introduction

Magnetic resonance spectroscopy (NMR) and imaging (MRI) have had a tremendous impact on research in physics, chemistry, biology and medicine. In chemistry and biology it has become the method of choice for structure analysis because of the wealth of information that can be obtained on a very local scale. For example it is possible to extract inter-atomic distances and coupling strengths within a molecule that can shed light on both the 3D configuration and functional behavior. In materials science NMR is the method of choice for analyzing materials with some inherent disorder present. In medicine MRI has become the most prominent diagnostic tool because of its versatility and non-invasive character. At present there are no adverse effects known of the static magnetic fields or the rf irradiation involved. This is mainly due to the low energy scales involved. Even at the highest possible fields the nuclear Zeeman splitting, and thus the electromagnetic radiation to pump the transitions, remains much smaller than the thermal energy $k_B T$ at room temperature. However this advantage also has an adverse effect. The main limitation in NMR and MRI is the fact that these low energy scales inevitably lead to a rather low sensitivity. Only with the utmost care in noise reduction can the spectra or images be accumulated in a reasonable amount of time. Even then, the technique is only applicable for samples with relatively large...
concentrations of spins. In spectroscopy typical sample sizes are of the order of 10–100 mm³ and in MRI one generally is very happy to obtain a 1 mm resolution. In contrast to various other techniques one typically requires of the order of 10⁴–10⁶ nuclei in the sample. In analytic chemistry, the sensitivity severely limits the application as a routine in situ analysis tool.

It is clear that there are many research topics that would benefit from even a modest improvement in sensitivity or resolution. An improvement in sensitivity by a factor of 10 would bring the data acquisition times down from days to minutes and would allow for example online quality control in chemical or pharmaceutical production. Also, if the MRI resolution could be boosted to the micron scale, it would be possible to perform functional studies on the level of individual cells.

So far these expectations are far from being realised and in fact there are firm physical laws that stand in the way of spectacular short term improvements. Nevertheless, many research groups are working towards gradual improvements and in retrospect we do have seen an improvement of nearly such a factor of 10 in the last decade, and there is no reason to be pessimistic about the future. It is therefore that we witness a renaissance of technique development in magnetic resonance striving to improve the sensitivity. For example there is a renewed interest in Dynamic Nuclear Polarization (DNP), the use of optically polarized Xe and ³He, and magnetic resonance striving to improve the sensitivity. For there has been an improvement of nearly such a factor of 10 in the last decade, and there is no reason to be pessimistic about the future. It is therefore that we witness a renaissance of technique development in magnetic resonance striving to improve the sensitivity. For example there is a renewed interest in Dynamic Nuclear Polarization (DNP), the use of optically polarized Xe and ³He, and magnetic resonance striving to improve the sensitivity.

First we will shortly point to recent reviews in the literature on both microcoils and force detected NMR microscopy. In the next paragraphs we will discuss the issues at hand on an elementary level and give some examples of recent developments in the field. Webb has reviewed the essentials of solenoid microcoils.1 Although this paper dates back to 1997, it contains most of the important aspects including application for liquid chromatography and electrophoresis. In a more recent paper by Ciobanu et al.2 the state of the art in MR spectroscopy is summarized with some emphasis on biological problems. The impact of rf field inhomogeneity with various imaging sequences is given with a discussion on sensitivity and resolution.

The aspect of rf field homogeneity and optimization of sensitivity for microcoils is analyzed in detail by Minard and Wind.3,4 These two papers contain many useful guidelines for the design of solenoid rf coils with specific optimization results for the number of windings and wire diameter for either conducting and non-conducting samples.

A theoretical analysis of the rf field homogeneity in high frequency solenoid coils is given by Engelke.5 The electromagnetic field is derived analytically and used in numerical computations to analyze the effect of dielectric sample materials, and standing waves that may occur if the geometrical sizes are not small compared to the wavelength of the rf radiation. Eroglu et al.6 reviewed the design and performance for spiral surface coils in particular for micro fluidic applications. In general, a good local sensitivity can be achieved at moderate spectral resolution. Integration in micro fluidic devices is indeed possible, although the close proximity of the rf coil degrades the field homogeneity (and thus spectral resolution) because of susceptibility problems. The same group also reviewed the various applications of these surface coils in particular for biological samples.7 A detailed description of the use of microcoils in micro fluidic devices is given by Masson et al.8 The achieved sensitivity for a test sample of sucrose in D₂O gives a proof of principle for 'lab on a chip' analytic capabilities. Note, however, that these tests are done at relatively high magnetic field values that require state of the art superconducting magnet systems. For routine applications it will be desirable to use permanent magnets, in order to make the spectrometer compact and robust for in situ applications in e.g. a production environment. At present it is not clear whether it is possible to achieve the required resolution in magnet systems that are not infinitely large compared to the size of the micro fluidic devices.

The implementation of microcoils for solid state NMR is given by Yamauchi et al.9 They emphasize the use of microcoils to generate very high rf fields that are needed to excite quadrupolar nuclei where the spectral distribution can be as broad as several MHz.

The limits of the attainable resolution in magnetic resonance spectroscopy are discussed in a recent review by Glover and Mansfield.10 They give a detailed historical overview of the developments in this field and re-examine the fundamental principles of imaging and signal detection to evaluate the limits in resolution, in particular for liquid, solid and gas-phase microscopy.

**NMR sensitivity and strategies for improvement**

The basic expression that gives the NMR signal to noise ratio is given by:

\[
\frac{S}{N} = \frac{k_B c_0}{\sqrt{\gamma N g^2 h^2 (1 + 1) k_B T^3 \lambda^2}} \sqrt{4 k_B T R_{one-M}}
\]

where \(k_B\) is a scaling factor accounting for the rf inhomogeneity of the coil, \(B_{1i}\) is the magnetic field induced in the rf coil per unit current, \(V_s\) the sample volume, \(N\) the number of spins per unit volume, \(\gamma\) the gyromagnetic ratio, \(I\) the spin quantum number, \(\omega_0\) the nuclear Larmor precession frequency, \(T\) the temperature and \(h\) and \(k_B\) are Planck’s and Boltzmann’s constant, respectively. The denominator describes the noise using the noise factor of the spectrometer (\(F\)), conductive losses of the coil, circuit and sample (\(R_{one-M}\)), and the spectral bandwidth (\(\lambda\)). In the derivation of this expression it is assumed that the noise figures are dominated by the coil resistance, including sample losses. An implicit assumption is that the noise of the preamplifier is negligible. In this case, the Q-factor of the resonant LC circuit drops out of the problem since signal and noise are both amplified by the same amount. This assumption may not be true any more if the noise is suppressed as for example in cryo-cooled detector coils. From the above equation it is clear that most of the parameters are constants of nature or dictated by the sample properties or sample conditions such as temperature. The basic strategy to optimize the NMR signal is to choose the lowest possible temperature combined with the highest magnetic field (and thus rf frequency) that is available. It has been widely accepted that a cooling of the rf coil in so-called cryoprobes is a sensible way to improve sensitivity. For pure metals, the resistivity is more or less linear in temperature, so a reduction in temperature by a factor of ten may lead to a factor of about three reduction in the noise amplitude. Commercial instruments now provide up to 900 MHz proton frequency in a 21 T static field. Alternatively one can go to a resistive magnet or combined hybrid magnet up to 45 T in one of the (inter)national high field magnet facilities. Next, one wants to optimize the filling factor of the specific rf coil, or maximize the \(V_s N\) product. The only technical design parameters of the rf coil and detection that can be manipulated is in fact the term \(B_{1i}\). In other words, one should optimize the rf field that is generated per unit current. In addition, one should minimize the total losses in the rf circuit to reduce the noise. This translates to the question of choosing the proper coil geometry for any given sample. Minard and Wind3,4 have summarized the basic guidelines for the most common coil geometries. Note that for a uniform rf field the length of the wire that forms the coil must remain short compared to the wavelength of the rf frequency. For the highest frequencies this strongly
restricts the number of windings in the coil. Also, most metals are weakly diamagnetic and for example the copper wire that forms the coil will lead to a (weak) distortion in the magnetic field and thus will reduce the spectral resolution that can be obtained. In solid state NMR the resonance lines are rather broad because interactions do not average out as in solutions. For solid state samples the requirement for the field homogeneity is therefore less stringent and the coil can in general be designed to fit tightly around the sample holder.

In the following we will describe some rather unconventional approaches to achieve the best sensitivity for very small samples. For these mass-limited samples the expected signals will be very weak and one is forced to optimize the detection sensitivity as good as possible. As a general design concept one can formulate the following guidelines:

The volume in which the rf field is generated should be confined in space to maximize the energy stored in the coil. This is equivalent to minimization of the self inductance $L$. In this space the rf field should be homogeneous and for optimum performance should be completely filled with sample material. In the example of a traditional helical solenoid roughly half of the magnetic field energy is outside the coil and is wasted. Also, the field is linearly polarized and for a circular polarized spin precession only half of the signal is detected. Thus the effective efficiency of a helix is only about 25% of the theoretical limit.

The resistive losses should be minimized. This translates to a geometry where the surface current density is a constant and covers (nearly) the entire surface of the sample volume. Eddy currents should be avoided where possible. For this point, a conventional helix is a good choice. Small improvements can be made by using rectangular cross section wire, although the distance between the windings should not be too small to prevent arcing when used at high power levels.

Depending on the intended frequency range, one should be careful to keep the total path length of the coil much smaller than the rf wavelength. If this condition is not met, the self-capacitance of the coil may lead to a current density that varies along the coil axis, severely reducing the rf homogeneity. Also, one should realize that the spins at the far end of the coil are excited at a later time than the top end. During the detection cycle the currents should have the same direction (phase). For example in a non resonant setup with a thread that is cut on a CNC lathe. In a later version of this design a variable pitch in the coil to get optimum rf excitation fields even at moderate power levels. A final point of attention should be toward a minimization of the length of the electrical leads between capacitor and inductor since this all adds to the electrical losses. Although the length of wiring is strongly reduced by scaling down the size of the helix, the effective resistance is nearly constant because the diameter of the wire is scaled by the same amount. Note that the rf currents flow only in a surface layer given by the penetration depth near the inside of the helix. The field of a cylindrical coil is at a maximum at the center and given by the simple expression:

$$B_1 = \frac{\mu_0 n}{2r} \left[ 1 + \left( \frac{j}{2r} \right)^2 \right]$$

where $\mu_0$ is the permeability of free space, $n$ the number of turns in the coil, and $r$ the radius and length of the coil. This shows that the sensitivity increases with the inverse of the diameter for coils with a fixed form factor $f = 2r/n$, which is the basic rationale for using microcoils to get NMR spectra from mass/volume-limited samples.

The $B_1$ field falls off to about half the center field near the ends of the helix. This inherent inhomogeneity of the rf field is further amplified by the fact that parallel currents repel each other, leading to a redistribution of the current away from the axis for the outermost windings. A final cause of rf field inhomogeneity is related to the current leads that break the cylindrical symmetry. A heat design to minimize most of these problems is sketched below in Fig. 1. Here, the helix is embedded in the center of a custom-made capacitor disk, which provides an LC geometry with very short connections. Since the currents in the capacitor plates are essentially radial, a minimum distortion of the rf field pattern is achieved. The microcoil itself is wound on a vessel coil former with a thread that is cut on a CNC lathe. In a later version of this design we introduced a variable pitch in the coil to get optimum rf homogeneity in the center half of the sample space (see Fig. 2 and 3). In principle this is similar to the classical trick of so-called “end-compensated” rf coils in which the outermost windings were squeezed to a slightly smaller pitch. A reduction to below 100 micron diameter is possible but the machining and handling of such coils will be rather tedious. For this reason various groups have

**Solenoid microcoils**

The classical geometry to create a magnetic field with an electrical current is the solenoid coil or helix. Even for a limited number of windings this geometry provides a reasonable homogeneous $B_1$ field and a good filling factor is possible by winding the coil directly onto a holder containing the sample. Miniaturization to a scale of several hundred microns is not very difficult although the wire diameter (typically 20 to 50 micron) becomes very small and a freestanding coil is a very delicate object. For classical LC resonant circuits one is tempted to use as many windings as possible to get reasonable inductance values (of the order of 10–20 nH). However the penalty one pays is that the maximum rf power that is allowed without severely heating the sample may be rather low. Fortunately a good $B_1$ field factor helps both to enhance the detection sensitivity and to get high rf excitation fields even at moderate power levels. A final point of attention should be toward a minimization of the length of the electrical leads between capacitor and inductor since this all adds to the electrical losses. Although the length of wiring is strongly reduced by scaling down the size of the helix, the effective resistance is nearly constant because the diameter of the wire is scaled by the same amount. Note that the rf currents flow only in a surface layer given by the penetration depth near the inside of the helix. The field of a cylindrical coil is at a maximum at the center and given by the simple expression:

$$B_1 = \frac{\mu_0 n}{2r} \left[ 1 + \left( \frac{j}{2r} \right)^2 \right]$$

where $\mu_0$ is the permeability of free space, $n$ the number of turns in the coil, and $r$ the radius and length of the coil. This shows that the sensitivity increases with the inverse of the diameter for coils with a fixed form factor $f = 2r/n$, which is the basic rationale for using microcoils to get NMR spectra from mass/volume-limited samples.

The $B_1$ field falls off to about half the center field near the ends of the helix. This inherent inhomogeneity of the rf field is further amplified by the fact that parallel currents repel each other, leading to a redistribution of the current away from the axis for the outermost windings. A final cause of rf field inhomogeneity is related to the current leads that break the cylindrical symmetry. A heat design to minimize most of these problems is sketched below in Fig. 1. Here, the helix is embedded in the center of a custom-made capacitor disk, which provides an LC geometry with very short connections. Since the currents in the capacitor plates are essentially radial, a minimum distortion of the rf field pattern is achieved. The microcoil itself is wound on a vessel coil former with a thread that is cut on a CNC lathe. In a later version of this design we introduced a variable pitch in the coil to get optimum rf homogeneity in the center half of the sample space (see Fig. 2 and 3). In principle this is similar to the classical trick of so-called “end-compensated” rf coils in which the outermost windings were squeezed to a slightly smaller pitch. A reduction to below 100 micron diameter is possible but the machining and handling of such coils will be rather tedious. For this reason various groups have
explored the option to use planar lithographic methods to define the coil shape.

For solenoid coils adding more turns to the coil will enhance the $B_i/i$ ratio and thus both the inductance and the signal response. At the same time the coil resistance will increase linearly, so the improvement in sensitivity will be proportional to the square root of the number of turns $n$. Note, however, that as soon as a packing of one is reached, i.e. adjacent turns are in close contact, one can only add more turns in a single layer if the wire diameter is reduced. This will give a nearly quadratic increase in resistance as a function of $n$, so the sensitivity will become independent of the number of turns. At the same time we will have a larger ohmic heating at the center of the coil and an enhanced danger for arcing, so the optimum is generally found for only a limited number of turns.

Besides rf performance, static field distortions due to susceptibility effects are an important factor in the design of microcoil probeheads. The close proximity of the coil wires to most of the sample will adversely affect the spectral resolution. As was described by Webb,1 possible workarounds are the immersion of the coil in a material with a susceptibility similar to that of the coil. In this way one mimics an infinite cylinder of given susceptibility in which the static field will be homogeneous. Alternatively the coil can be designed of multiple layers of different metals (e.g. aluminium and copper) with compensating susceptibilities. In solids samples magic angle spinning can (partly) average susceptibility effects.

**Planar microcoils**

The most common geometry used in planar microcoils is based on a spiral design with the center winding contacted to the outside using a connection to another layer which is electrically isolated with a thin oxide layer. A typical coil layout with the corresponding rf field profile is shown in Fig. 4. Using standard lithography techniques as used in the micro-electronics industry one can define any structure with dimensions down to below a micron. Also the fact that mass production is possible without the tedious procedure of winding a wire onto a coil former is quite appealing. Nevertheless, a detailed analysis shows that such a surface spiral has some serious drawbacks, compared to a helical coil. First, the outermost windings are much less efficient in the sense that they contribute less to the centre axial field while they largely dominate the resistive losses. Second, the products fields by the outer windings cause considerable eddy currents in the centre windings adding additional losses and lowering the field homogeneity in the centre region. For this reason the optimum is found for only a few windings, and thus a rather low inductance. A third less advantageous aspect of surface coils is that the conductance of the connection to the bottom layer is not ideal. Using typical thin film methods as is common in micro-electronics one ends up with additional losses that can be comparable to the losses of the coil itself.

One of the appealing aspects of lithographically produced rf coils is that it may fulfill the promise of a ‘lab on a chip’ in which the material synthesis and analysis is integrated on a single chip. An example of such a probe is shown in Fig. 5. It is even conceivable to integrate the NMR rf source and data acquisition electronics on the same chip. The fast developments and miniaturization in the cell-phone industry (working at rf frequencies in the same band as employed in NMR) show that this is not unrealistic. In fact, a design integrating coil and preamplifier module was developed using CMOS technology by Boero et al.12

A further possibility would be the use of microcoils for performing solid-state NMR spectroscopy in very high field levels.
resistive magnets. In general these magnets are designed for general purpose physical experiments and lack the homogeneity needed for high resolution NMR. The restricted volume of the microcoil environment might alleviate the homogeneity issue and one should be able to perform NMR spectroscopy on e.g. a 100 micron single crystal with acceptable resolution.

**Intermezzo: quadrupolar nuclei**

As is clear from the periodic table of the elements shown in Fig. 6, the majority of the nuclear isotopes have a quadrupolar moment. This means they are encountered in many relevant materials such as minerals, structural ceramics, semi-conductors, glasses, catalysts etc. For quadrupolar nuclei in natural and synthetic samples one is often faced with the problem that resonance frequencies are dispersed over several MHz due to the quadrupolar interaction. The quadrupolar interaction is a valuable source of structural information but compromises spectral resolution or even NMR detectability. Furthermore, it can be a problem to obtain crystals of sufficient size to undertake a single crystal study of a material. Application of microcoils may help to solve both problems. This may be relevant for example in isotopically enriched materials or archeological specimen of minute sizes. Also in the case of zeolites that are used in catalysis only very small single crystals are available. As an illustration the $^{27}$Al NMR signals of a 30.4 μg single crystal of sapphire were observed at 14.1 T in a 300 μm diameter solenoid microcoil. Sapphire has a single Al site with a quadrupolar coupling constant of 2.4 MHz and $\eta = 0$. Fig. 7 displays the spectrum of the single crystal obtained with a 150 ns pulse using an rf field of 1.7 MHz. All five single-quantum transitions ($5/2 \rightarrow 3/2, 3/2 \rightarrow 1/2, 1/2 \rightarrow -1/2, -1/2 \rightarrow -3/2, -3/2 \rightarrow -5/2$) are clearly visible in the spectrum displaying a quadrupolar splitting of 80 kHz. Because of the even excitation of the various lines, they are observed with intensities very close to the theoretical ratios for a non-selective excitation of a spin $5/2$ system (5:8:9:8:5). Using lower rf fields the intensities of the satellite transitions are often diminished because they are less efficiently excited. This is a particular problem for powder spectra were the intensity in the satellites is dispersed over a large frequency range and are often hardly observed with respect to the narrow central transition which is only broadened by the quadrupolar interaction in second-order. Fig. 8 shows the experimental and theoretical $^{27}$Al spectrum of powdered sapphire using a two-pulse echo. The powder lines of the satellite transitions are spread out over 720 kHz but are clearly observed without significant distortions, reflecting the features of a site with $C_{qcc} = 2.4$ MHz and $\eta = 0$ in accord with the literature data.

**Saddle coil and stripline geometries**

In MRI the large volumes and axial access requirements for the patient make a helical coil impractical and the use of saddle-coils is more or less standard. Also in high resolution liquid NMR this turns out to be the rf geometry of choice. The only requirement for NMR excitation and detection is that a sufficiently strong $B_1$ component is generated perpendicular to the static $B_0$ field. Clearly, a helical coil is by no means the only configuration that does the trick, and despite conventional wisdom it may not even be the most efficient configuration. Let us take for example a single wire of diameter $d$ along the $B_0$ axis, carrying an rf current $i$. Because of the finite penetration depth, the current will flow in a thin surface layer. The field produced by such a current has a cylindrical symmetry, oriented tangential to the wire and is given by Ampère’s law. At all positions in space this field is oriented perpendicular to the external static field and can be used to excite an NMR transition. For example a $\pi/2$ pulse for a sample shell at distance $r$ from the wire center will tip the nuclear spins from the $z$ to the radial direction in the $xy$ plane. By reciprocity the same wire can be used to detect the...
NMR signal. Even though the spins do not have a uniform precession in space they all contribute in-phase to the induction voltage across the wire. It is clear that such a simple wire is not a very efficient configuration, mostly because the rf field homogeneity is rather poor. To some extent this can be overcome by using a coaxial arrangement with an inner and outer conductor as shown in Fig. 9. Also the bird cage or saddle coil geometries (see Fig. 10) used in MRI are in essence just a collection of straight wires optimized to give the best rf field homogeneity combined with possible phased-array methods to get the best matching to a specific spin assembly.

As an instructive example let us consider a thin sample layer on a glass substrate. In a cylindrical helix one would have to cut up the sample in bars and stack a few of them on top of each other to fill the helix. An alternative would be to use the surface coil described earlier (Fig. 4). In the usual configuration with the axis of the coil oriented perpendicular to the external static field one would detect mainly the centre part of the sample located within the inner winding. This is illustrated in the top graph in Fig. 11 showing the proton signal of a 60 micron thick layer of cellulose acetate. A strong signal is obtained in only 16 scans. On the horizontal axis a series of spectra, taken with stepwise increasing rf pulselength, connected side by side. This type of nutation experiment can be used to determine the strength and homogeneity of the rf field. In this case, the signal intensity as a function of pulse length reveals a strong rf field of the order of 500 kHz at a power of only 5 W.

One should realize that on top of the windings the field is not oriented along the axis, but has a substantial radial component as well. This component can be used for NMR excitation and detection if we flip the surface coil by 90 degrees, such that the coil axis is now parallel to the static field axis. As in the example of the straight wire, the radial rf field does not give a uniform nuclear moment. In fact, at opposite points across the symmetry axis, the nuclear spins press out of phase and no intensity is lost. This becomes clear from inspecting the bottom graph of Fig. 11 showing the signal of the same cellulose acetate film but with the coil and sample rotated over 90°. Indeed a strong signal is obtained, with only slightly lower S/N and effective rf field strength than with the coil in the upright position. A bonus in this case is that although the B1i field factor is lower for the radial component, the effective sample area can be made much larger in order to yield a net improvement of the signal to noise ratio for a given sample thickness.

If we now concentrate on a small area containing n windings on top of the 2D helix it is only a small step to realize that the same field profile can be generated by a linear array of wires connected in parallel, in other words by a simple grid structure or single strip as shown in Fig. 12. The only difference is in fact the impedance seen by the external source. In order to generate the same B1 field we have to supply n times the current while both the inductively stored energy and the resistive losses remain the same. Although the signal amplitude is now reduced by a factor n, the same is true for the total noise in the parallel circuit and as long as the preamplifier noise can be neglected the performance remains the same.

In comparison with the planar helix we now have the advantage that the geometry does not have a topological pole, which means that all connections can be made in the same plane, removing the problems associated with the connection to another plane. Since the structure no longer has a cylindrical symmetry, the current
distribution is more homogeneous, and Eddy currents are avoided. Also, the surface can be scaled in a simple and trivial way for specific sample sizes. The main applications are anticipated to be in the field of thin coatings or epitaxial layers, Langmuir Blodgett films, phospholipid bilayers, membranes containing polypeptides and proteins etc.

As usual, however, there is no such thing as a free lunch. The low inductance of such a circuit makes it troublesome to implement the usual LC resonator configuration. Since the effective impedance is very low (both inductive and resistive) the most natural electrical configuration is to incorporate the coil in a tuned circuit like a quarter lambda transmission line where the strip is mounted as a short. A more serious problem is that the rf magnetic fields fill a large volume in space. Also the rf currents will try to avoid each other, leading to an enhanced current density at the edges of the grid or strip and thus to extra dissipation and lower $B_i$ field homogeneity. The advantages and disadvantages of the different coil designs discussed so far are summarized in Table 1.

A final point of concern will be the $B_i$ inhomogeneity introduced by the planar strips. This would be especially important for high-resolution liquid NMR. On the other hand the field produced by planar films is nearly homogeneous near the center of the strip and for samples away from the edges the effect should be small. A possible strategy to counter possible resolution problems is to use multi layer strips of "zero-susceptibility wire".

**Micro imaging**

Magnetic Resonance Imaging has become an indispensable diagnostic tool in hospitals around the world, despite the relatively low resolution, which is of the order of 1 mm in the most common instruments. With some effort a resolution of the order of 100 micron can be achieved in specialized scanners for small animals. In the ideal world one would like to be able to zoom in to the single cell or beyond to study the functional behavior of for example single neurons. However, the acquisition times would become excessive and $B_i$ single neurons. However, the acquisition times would become excessive and in vivo measurements are no longer possible. In patient diagnosis there has been a development of specialized local detection coils that are anatomically formed to obtain the best filling factor for a specific part of the body, such as the head or knee. In this paper we will only consider the concepts that can be applied in moving towards submicron resolutions.

The first prerequisite is to optimize the detection sensitivity to be able to measure the signals from small sample volumes or voxels. This inevitably leads to microcoil configurations as described above. In most of the experimental configurations a 2D symmetry was chosen, which simplifies the analysis and strongly reduces the measurement time. In-plane spatial resolutions down to one micron have been demonstrated for 2D phantoms with a slab thickness of 75 micron. In biological systems the diffusion limits the attainable resolution to a typical value of 10 micron. This relaxes the sensitivity requirements to a voxel of 10 µm³ with a total measurement time of an hour or less to avoid degradation of the samples. With the present state of the art detection hardware this is now possible. Note however, that this sensitivity is only quoted for protons in an aqueous environment, corresponding to a proton concentration of 111 M.

It is clear that an increase in spatial resolution can only be obtained if the magnetic field gradients are matched to this smaller length scale. In clinical MRI setups one typically uses switched gradients of the order of 1 G cm⁻¹ to achieve a resolution of about 1 mm. The field gradients in microcoil MRI will have to be of the order of 0.1 to 1 T cm⁻¹ to obtain submicron resolution. This can be achieved relatively straightforward by integrating the gradient coils on the same chip. In fact, the electrical power needed to drive the gradient coils can be very modest, and prototype 3D micro-imaging instruments have indeed demonstrated 0.5 T cm⁻¹ gradients.

A recent review of the developments in this field was given by Ciobanu et al., summarizing the rapid developments in MR microscopy. Fig. 13 shows a typical high resolution MRI image of a single oocyte cell. With the experimental resolution of about 10 micron one can clearly distinguish for example the cell nucleus. An example for local spectroscopy is shown in Fig. 14, producing a spectrum from a single neuron, containing both the nucleus and cytoplasm. The localized volume was about 220 × 220 × 220 µm³. As was stated in the introduction, there are firm physical laws.
that stand in the way of sensitivity improvements by several orders of magnitude. Therefore, if such improvements are desired, one has to turn to fundamentally different means of enhancing the spin polarization or come up with new approaches for detecting the magnetic resonance signal. In the following we will concentrate on the specific example of Force Detected Magnetic Resonance (MRFM).

Magnetic resonance force microscopy

A relatively recent method for improving the detection sensitivity of NMR is based on mechanical detection. Although the ideas with respect to the mechanical detection of the magnetic resonance phenomenon were proposed as early as 1964 and its concepts proven in 1967 for electron spins, mechanical detection of the nuclear magnetic moment was only achieved much later. The first method that was successfully applied to this end is called Magnetic Resonance Force Microscopy (MRFM) as proposed by Sidles and co-workers in 1991.\textsuperscript{14,15} It was demonstrated for magnetic resonance detection of electrons by Sidles in 1992\textsuperscript{14} and for protons by Rugar and co-workers in 1994.\textsuperscript{16} It uses a mechanical cantilever as is known from Atomic Force Microscopy (AFM) to detect forces exerted on a spin system in a very inhomogeneous magnetic field. A typical configuration for MRFM measurements is shown in Fig. 15. The development of this method has been driven mainly by the prospect that it may be possible to detect a single (electron) spin. This would not only be a spectacular demonstration in itself, it could also help to understand the quantum mechanical foundations and become an important tool in quantum computation.

Sidles \textit{et al.} gave an extensive review of the field of MRFM in 1995\textsuperscript{17} including both the design principles and potential applications. These design concepts are still relevant today. Until recently most groups in this field have concentrated on optimizing the force sensitivity of the MRFM cantilever. Only recently have groups shifted their attention to the development of the appropriate NMR sequences to be useful in imaging. The major drawback of the method in the context of NMR remains that the inhomogeneity of the field prevents any kind of spectroscopy. It has been suggested by Leskowitz \textit{et al.}\textsuperscript{18} that mechanical detection of NMR can be performed without the presence of the magnetic field gradient. A large scale model was demonstrated, but miniaturization to obtain a practical instrument proves to be difficult.

In the following we will describe the methods used in MRFM at an elementary level and compare the sensitivity with that using inductive detection. The basic mechanism used in MRFM is the fact that the nuclear magnetic dipole moment experiences a force when placed in an external gradient field. This force is detected by mechanical means using an atomic force microscope cantilever. The deflection of the cantilever can be measured very accurately with sub-angstrom resolution by optical methods. The force \( F \) acting on a magnetic dipole in an inhomogeneous magnetic field is given by:

\[
F = V \frac{\partial}{\partial y} (MB) B^y
\]

in which \( V \) is the selected sample volume, \( M \) is the sample magnetization and \( \nabla B \) is the gradient of the magnetic flux density. The MRFM experiment involves modulating the nuclear magnetization \( M \) by means of rf pulses and/or sweeps and detecting the associated change in the force. For a single \( m \rightarrow m \pm 1 \) transition the change in observed z-magnetization is:

\[
\Delta M_z = \frac{N g \mu B_0}{(2I + 1) k_B T} B_y
\]

Note that the observable magnetization change of the ensemble average is directly proportional to the square of the induced magnetic dipole \( \langle \Phi \rangle \) whereas the observable change in the magnetization of a single spin is proportional to the magnetic dipole \( \Phi \) itself and is independent of the ambient temperature and spin quantum number \( I \).

Since detection of the static forces exerted on the cantilever by the nuclear magnetic moment is very difficult, modulation of the nuclear magnetization is usually performed. There are several methods in use to vary the induced magnetic dipole by rf irradiation. The most common method is based on fast adiabatic passage (see Fig. 16). The rf field is placed very far off-resonance and repeatedly (dictated by the cantilever frequency) swept through the transition in an adiabatic fashion. The process will be adiabatic if the sweep rate is sufficiently slow so that the spin system will remain in thermal equilibrium. At the start of the experiment the magnetization is assumed to be in thermal equilibrium and aligned with the external field. Then the rf field is swept adiabatically
through resonance in order to obtain complete inversion of the (fictitious) $I = 1/2$ transition after which the rf field is swept back, approximately realigning the magnetization with the axis of thermal equilibrium. This process is repeated several times, until the spin-locked magnetization is completely decayed. This method is best used for materials with a long relaxation time $T_1$.

In the most common situation where both the field gradient and the modulated component of the magnetic moment are pointed along the $z$ axis, we can write the time dependent force $F(t)$ on the sample as:

$$F(t) = \int M(z,t) \frac{\partial B}{\partial z} dV $$

Furthermore, as a result of the presence of the magnetic field gradient the Larmor resonance condition, $\omega_0 = \gamma B_0$, varies over the sample, i.e. becomes spatially dependent. Thus we have the option to selectively excite slices from the sample through variation of the irradiation frequency or by altering the position of the magnetic field gradient source. In nuclear MRFM the gradient field is usually brought about by introducing a small magnetic particle in an otherwise homogeneous magnetic field ($B_0$). Since this external $B_0$ magnetic field is very strong the magnetic particle will be completely saturated by the field. The magnetic field at a distance $r$ from this magnetized particle depends on its actual shape.

The geometry that gives the strongest gradient for a given ferromagnetic material is that of a cone with top angle with $\theta = \sqrt{2}$, resulting in a field gradient:

$$\frac{\partial B(z)}{\partial z} = \frac{B_{sat}}{3\sqrt{2}z}$$

where $z$ is the distance from the top of the cone. For practical purposes a cylindrical shape is easier to handle and the resulting excitation slices are easier to analyse and to deconvolute into a real space image. The final goal for a scanning MRFM will be realized if the gradient is created by a ferromagnetic particle on the cantilever. In this case the weight limitations will dictate a spherical shape for the gradient source.

In case of a spherical ferromagnetic particle the gradient field is given by:

$$\frac{\partial B(z)}{\partial z} = -2B_{sat}\frac{R^2}{z^3}, \text{ for } z \geq R$$

where $R$ is the radius of the sphere and $z$ is the distance from the centre.

As can be seen from this expression the gradient field generated by a spherical particle is far from linear. This means that this type of gradient is less suitable for large distance sub-surface imaging since the gradient strength rapidly drops.

A cantilever is for most purposes well described as a mass-spring system with a resonance frequency $\omega_0$:

$$\omega_0 = 2\pi f_0 = \frac{k_s}{m}$$

where $k_s$ is the spring constant and $m$ is the mass of the cantilever plus sample. The mechanical oscillator has a quality factor $Q_c$:

$$Q_c = \frac{k_s m}{b}$$

where $b$ is the friction in the cantilever beam.

As in the LC resonator, the detection limit is dictated by the noise in the system. In a mechanical resonator the noise is due to the thermal excitation of vibrations at the resonance frequency. If we set the force detection limit in a measurement bandwidth of $\Delta \nu$ to a SNR of 1 then the smallest detectable force can be given by:

$$F_{min} = \frac{4k_B T \Delta \nu}{Q \omega_0}$$

Combining these equations leads to the signal to noise ratio of the MRFM experiment:

$$\text{SNR} = \frac{F}{F_{min}} = \frac{\partial B}{\partial \xi} \frac{Q \omega_0}{\alpha} \frac{\Delta \nu}{\sqrt{4k_B T \Delta \nu}}$$

If we compare the expressions for the SNR in the mechanical and inductive detection methods for the nuclear magnetization assuming equal detection bandwidth we find:

$$\frac{\text{SNR}_{mech}}{\text{SNR}_{inductive}} = \frac{\frac{\partial B}{\partial \xi} \frac{Q \omega_0}{\alpha}}{\frac{\partial B}{\partial \xi} \frac{\Delta \nu}{\sqrt{4k_B T \Delta \nu}}} = \frac{Q \omega_0}{\sqrt{4k_B T \Delta \nu}}$$

From this equation we can immediately see what are the strong points of mechanical versus inductive detection are: First, the inductive signal scales with $B_1$ (per unit current). Numerically, this is typically 5 orders of magnitude smaller than the gradients used in MRFM (~10 mT versus several kT m$^{-1}$). Also the losses in mechanical resonators can be very small, leading to a typical $Q$ of 10$^5$ versus 10$^2$ in a typical LC circuit. The numerical value for the square root of the ratio $R/B$ is then typically of the order of 10$^3$. If we scale the two methods to smaller sample sizes, then the MRFM method gains in sensitivity mainly because larger gradients can be used. The scaling for inductive detection gives only a moderate increase in $B_1$ while the losses remain essentially the same. The only factor working in favour of the inductive method is the Larmor frequency in the denominator. This is simply a consequence of the fact that the induction voltage scales with the time derivative of the total magnetic flux in the coil. For protons a break-even point is typically reached for samples of a few hundred microns in size. In smaller samples the MRFM method will be more sensitive. For a 10 micron sample in a hypothetical rf coil of 50 micron diameter, the inductive detection limit will be typically of the order of 10$^{12}$ spins in a bandwidth of 1 Hz, corresponding to a concentration of about 1 mol L$^{-1}$. For the same sample size, a mechanical detection may give about two orders of magnitude better sensitivity. An example for MRFM imaging in very high field gradients is given in Fig. 17, showing an effective resolution below 1 micron.

Note again that these numbers should be treated with care. The most common samples in MRFM will be in the solid state where the static line width is seldom in the range of a few Hz. Common tools such as Magic Angle Spinning or special NMR pulse sequences are not applicable. Also we should stress again that high resolution spectroscopy in the presence of very large static field gradients is not possible. A special niche application could be the imaging for low $\gamma$ nuclei in relatively low fields generated by permanent magnets. Because of the intrinsically large static gradients involved one is not limited by the lower homogeneity of these magnet systems. Mechanical detection will not be competitive for low concentration samples, and for the time being is mainly relevant for micro imaging purposes. The use for spectroscopic analysis will be limited, unless a method for zero gradient detection becomes available.

The highest sensitivity to date has been demonstrated by the group of Sidles$^{19}$ where they used a field gradient of 0.25 MT m$^{-1}$ with a sample of only 6.3 ng. The demonstrated detection sensitivity corresponds to about 75 Bohr magnetons in a one Hertz bandwidth. They propose that a new generation MRFM with a very tiny cantilever (3 micron long, 30 nm diameter) it should be possible to reach a force sensitivity of $10^{-20}$ N V Hz$^{-1}$ which is sufficient for single (electron) spin detection.

Nestle et al.$^{20}$ have evaluated the applicability of the MRFM technique for chemical investigations using either relaxation times or chemical shifts. They describe the various NMR approaches for

\[ F_{min} = \frac{4k_B T \Delta \nu}{Q \omega_0} \]

\[ \text{SNR} = \frac{F}{F_{min}} = \frac{\partial B}{\partial \xi} \frac{Q \omega_0}{\alpha} \frac{\Delta \nu}{\sqrt{4k_B T \Delta \nu}} \]

\[ \frac{\text{SNR}_{mech}}{\text{SNR}_{inductive}} = \frac{\frac{\partial B}{\partial \xi} \frac{Q \omega_0}{\alpha}}{\frac{\partial B}{\partial \xi} \frac{\Delta \nu}{\sqrt{4k_B T \Delta \nu}}} = \frac{Q \omega_0}{\sqrt{4k_B T \Delta \nu}} \]
creating a modulated spin magnetization that can be measured with force detection methods. From these experimental observables one can deduce the actual material parameters like the spin–lattice relaxation time $T_1$, the chemical shift and the diffusion properties. They discuss in particular applications in the field of coatings, colloids and nano structured semiconductors.

A first example of high resolution imaging with a structural contrast has been demonstrated by exploiting the quadrupolar interaction. The response of quadrupolar nuclei to an on-resonance pulse depends on the ratio of the quadrupolar interaction and the rf field strength. As the external field gradient does not play a role, this nutation behavior can be exploited for imaging materials with a contrast function depending on the local structure experienced by the nuclei. A pulse acting as a π/2 pulse on spins in a cubic environment (\(\alpha_0 = 0\)) will be felt as a π– pulse by spins in a distorted environment (\(\alpha_0 \geq \alpha_0\)). Applying such a “quadrupolar filter” in an imaging experiment allows one to enhance or suppress signals of nuclei in a specific position in the overall image. This is demonstrated in Fig. 18, showing the 1D image of a NaCl particle on top of a Na$_2$C$_2$O$_4$ crystallite. As NaCl has a cubic structure, 23Na does not experience a quadrupolar interaction, whereas in sodium oxalate the quadrupolar interaction is very large (\(Q_{ox} = 2.5 \text{ MHz}\)). The left image is the original MRFM image reflecting the 23Na spin density with an 11 μm resolution. In the right image a quadrupolar filter was applied prior to the imaging experiment allowing one to practically eradicate the signal from the cubically symmetric NaCl particle from the image; the contrast of the oxalate is only slightly affected due to offset effects during the nutation filter pulse.21

Let us assume a force sensitivity of the cantilever with optical interferometer of $5 \times 10^{-16} \text{ V Hz}^{-1}$. (This sensitivity was reported at 300 K in 1994 by Rugar et al.20). In a gradient of $\sim 10^5 \text{ T m}^{-1}$ this corresponds to a moment sensitivity equal to $\sim 10^{10}$ protons (300 K, 14 T). For a crystal of $\sim 10 \mu m$, a depth resolution of $\sim 1 \text{ nm}$ should be attainable for samples abundant in protons (lower gamma nuclei and lower concentrations lead to correspondingly lower resolution). This calls for a design with the appropriate mechanical stability to match this sensitivity.

In imaging one should realize that on the timescale of a typical NMR (or MRFM) experiment the real or spin diffusion can be substantial, therefore a resolution far below 10 micron will be difficult to achieve in biological samples. On the other hand, the local nature of the MRFM experiment allows a study of spin–diffusion in detail and provides an alternative route to get functional contrast. Also the very strong local field gradients used in MRFM can even help to suppress spin diffusion.22

A method for Fourier encoding of the spatial distributions of spins was proposed by Kempf and Marohn.23 In this method a magnetic gradient is shuttled in synchrony with the rf pulse sequence so that the local precession rate becomes proportional to the z coordinate. In this case a real-space reconstruction of the spin density map can be obtained by a linear Fourier transform of the integrated force signals.

**Conclusions and outlook**

Recent developments in magnetic resonance as described above show the liveliness of this field even after more than 50 years of its first inception. To date, high-resolution NMR is the method of choice in a wide variety of applications ranging from simple analytical compound identification to sophisticated 3D structure elucidation of biomacromolecules. Structural genomics and metabolomics are new fields of research that heavily rely on NMR information. Especially for the latter case, development of adequate microcoil technology is expected to find its way into this area of research e.g. in the form of integrated microfluidic devices. We furthermore foresee a role for microcoils for in situ analysis in production facilities using ex situ magnet assemblies. In combination with chemometric techniques for data-analysis the relatively low-resolution information obtained this way might be valuable for analysis and control of e.g. food products.

MRI has revolutionized medical diagnostics, which, with a judicious implementation of microcoils can be brought down to the single cell level. The possibility to image e.g. individual firing neurons are an exciting prospect. Brain slices are an established method for examining brain tissue function, and constitute an important intermediate between studies of single cells and of the whole animal. Examination of such samples by means of (cooled) microcoils, allowing a sensitivity and spatial resolution previously unattainable, well below the thickness of a cortical layer (< 100 μm) should prove a breakthrough in functional imaging.

In materials science, solid-state NMR is the only method that can probe local order and dynamics independent of the physical state or crystallinity of the material. The level of sophistication of solid-state NMR techniques has advanced to such a state that detailed structural information is obtained from complex biological structures such as membrane-bound peptides. An important strategy in this field is to exploit the information that is contained in the anisotropic NMR interactions by using oriented samples. Sensitivity optimized surface rf structures should prove to be an enabling technology in this field as they provide the possibility to work with limited amounts of (isotopically labelled) samples. Furthermore, the high rf field strengths will make existing experiments more easy and allow the implementation of more advanced schemes currently prohibited by power requirements.

High rf fields should also prove to be invaluable in inorganic materials science where various quadrupolar nuclei are encountered. The study of these nuclei is often hampered by their excessive line widths. Using microcoils adequate rf fields can be generated to directly excite the whole spin system and allow spectroscopy involving multiple-quantum transitions or two-photon absorption.

Finally, mechanical detection of magnetic resonance constitutes a major breakthrough in magnetic resonance sensitivity. The detection of a single nuclear spin is a more than exciting prospect which might be achieved in the near future. The limits placed on the
implementation of MRFM are imposed by the relaxation behaviour of the spin systems under investigation and the possibility to adapt the detection setup to the specific requirements of the material under study. In relevant materials amenable to studies by MRFM half-integer quadrupolar nuclei are abundant. It is possible to perform spatially localized spectroscopy enabling one to superimpose a contrast function depending on the local lattice structure onto images. In this respect imaging with true chemical specificity has been achieved. In combination with future sensitivity enhancements bringing the resolution in the 10–100 nm range this should provide a very powerful local probing technique.

Acknowledgments
The authors would like to thank Mr J. van Os and Mr G. Janssen and Dr K. Yamauchi for support and advice. Mr B. van den Berg is acknowledged for his craftmanship in micromachining the micro-coil parts. This work is part of the research programme of the Stichting voor Fundamenteel Onderzoek der Materie (FOM, financially supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO)) and Philips Research. We thank Dr R. Dekker, Dr J. Jansen, Dr P. van Tilborg and Dr J. Pikkemaat (Philips Research) for their interest and support. Dr R. Dekker provided the helical surface coils.

References

Fig. 18 Example of a $^{23}$Na MRFM image (resolution 11 μm) where the quadrupolar interaction is used to distinguish between different chemical environments of the nuclei. As the response of a quadrupolar nucleus to an rf field depends on the ratio of the rf field strength and the quadrupolar interaction experienced by the nucleus, we can get localized information about structural aspects of a material. This can be used to add a contrast function based on local symmetry to images of materials, as is shown above for $^{23}$Na in a NaCl crystallite (cubic symmetry) on top of a Na$_2$C$_2$O$_4$ crystallite with Na in a distorted surrounding. The left panel shows the unfiltered image. In the right panel signals from Na in a symmetric surrounding are eliminated from the image. Reproduced with permission from ref. 21.