1: Introduction

The second-order soundfield microphone represents a potential significant advance to the "state of the art" in coincident microphone array technology; this thesis presents the first comprehensive theoretical treatment of such a microphone, providing analysis and information that will be necessary if it is to be developed as a practical recording tool.

The theory of the second-order soundfield microphone is closely related to the theory of the first-order soundfield microphone and of second-order ambisonic surround sound systems, as well as to the general theory of second-order pressure gradient microphones. Hence, this thesis also presents an account of the theory of the first-order soundfield microphone which is intended to "fill in the gaps" in the existing literature, extends the existing analysis of second-order ambisonic surround sound systems, and includes a systematic and unified account of the theory of first-order and second-order pressure gradient microphones which the author believes to be more satisfactory than the treatments generally available in the literature.

1.1: Historical & Contemporary Context

One of the characteristics of a sound which we are capable of perceiving is, in many cases, the location of its source. As far as music is concerned one can, when present at a live performance, often determine at least approximately the positions of individual instruments, and this is considered to contribute to the appreciation and enjoyment of the performance. The ability to distinguish the sound of particular instruments within a group, or to hear a soloist through an accompaniment, is partly dependent on being able to locate them spatially. Furthermore, in some forms of music the spatial distribution of sound sources is considered to have an explicit and significant artistic role [17] [31] [67].

The acoustics of the venue in which a performance takes place also contribute to the overall aesthetic impression. As well as direct sound from the instruments, the listener hears sound which has been reflected, possibly multiple times, from objects and surfaces within and surrounding the listening environment (usually, of course, including the walls, floor and ceiling). Such reflected sound may be perceived as discrete echoes or as reverberation. The fine temporal and spatial structure of reverberant sound is not usually consciously appreciated, but nevertheless it conveys information about both the environment and the location of the instruments; there is a complicated but specific relationship between the position of a sound source and the resulting reverberant sound. Thus, reverberation assists in the localisation of instruments, as well as contributing to the sense of "space" and envelopment in the sound experienced by the listener [10] [14] [31] [33].

It is clearly desirable, therefore, that a sound recording and reproduction system should be such as to capture and recreate the spatial information present in the original sound field, including both direct and reverberant sound, as accurately as possible. A system which achieves this can provide the listener with the same impression of the location of sound sources, and of the ambience of the venue, as he would obtain if actually present at the performance. Any system which is not capable of conveying spatial information can reasonably be said to impose a form of distortion, which may be termed "directional distortion", on the reproduced sound; high fidelity reproduction, of course, requires that all forms of distortion are minimised [15] [31] [32] [33] [41].

With many studio recordings, it is of course the case that no performance, in the conventional sense, ever occurs. In such a situation, the spatial information discussed above simply does not exist; a sound reproducing system should instead accurately convey to the listener the impression of an artificially created spatial arrangement of sound sources, or "sound stage". There exists considerable freedom to use the sound stage creatively; a "natural" effect, as if all the musicians had actually been present and performed together, may, but will not necessarily, be desired.

The material presented in this thesis is primarily concerned with the capture of directional information from an original sound field; while it will occasionally be appropriate to mention techniques for the creation of synthetic sound fields, these are essentially outside the scope of the reported work. The author wishes however to distance himself from the prejudice, regrettably prominent in the literature, that the existence of an actual acoustic event to be reproduced necessarily confers any artistic superiority.

Although it is common to discuss the spatial element of reproduced sound with reference to music, it should be remembered that the positions and movements of sound sources and the ambience of the environment are perhaps even more important in other contexts, which may or may not involve an original acoustic event and may or may not be associated with a visual image. Examples include broadcast radio (or direct-to-disc) drama, recordings of theatrical performances, film or television soundtracks, and even such esoteric “niche market” material as recordings of railway locomotives arriving at a London station [96]. The correct reproduction of spatial sound can also be of benefit in virtual reality and telepresence applications [59] [60].

Stereo (or “stereophonic”) and surround sound systems provide, or at least are intended to provide, the capability to record and reproduce the spatial aspects of sound. Most systems described as stereo generate a reproduced sound stage of limited spatial extent (typically covering an angular distance of approximately 60° in front of the listener), while surround sound systems are intended to extend this so that the listener is completely surrounded, either in the horizontal plane or in three-dimensional space, by directions from which sound may be heard [32]; however, this distinction is not well defined and is not always adhered to. Surround sound systems in which the reproduced sound stage is confined to the horizontal plane are termed “pantophonic”, while those which also include height information and can therefore provide full three-dimensional (or “full-sphere”) reproduction are called “periphonic” [31] [47].

There is an unfortunate tendency in modern usage to describe as stereo only those systems employing exactly two transmission channels and two loudspeakers; this is certainly not correct. Terminology such as “three-channel stereo” or “multi-channel stereo” is quite acceptable. The term “multi-channel” in such a context usually indicates that the system employs more than two transmission channels [32].

No system of sound reproduction employing only a small number of loudspeakers can physically create a sound field identical to the direct and reverberant sound generated by a multiplicity of sound sources located in arbitrary directions, so the directional effect created by such a system must in some sense be an illusion. The objective of surround sound technology is therefore to arrange for the loudspeaker outputs to be such that the illusion is as convincing as possible; this requires a knowledge of “directional psychoacoustics”, i.e., the relationship between the physical sound field and the subjective sensation of directionality received by the listener [45] [46] [83].

So far as the recording and reproduction of a live performance is concerned, it is clear that spatial information cannot be reproduced if it is not first correctly recorded. The first-order soundfield microphone is known to be an extremely effective tool for capturing the directional properties of a sound field. Such microphones are closely associated with so-called "ambisonic" surround sound systems, but are in no way restricted to use with these alone; they have proved to be of considerable utility for conventional two-channel stereo recording, and may also be used to make recordings intended for playback over other surround sound systems [20] [29] [94].

The second-order soundfield microphone would provide the capability to capture a greater quantity of directional information. Although suggestions that such a microphone could be constructed appeared almost three decades ago [31] [40], and its potential use in conjunction with ambisonic reproduction systems has been mentioned from time to time [5] [31], no serious consideration has yet been given to its development, and consequently no comprehensive account of the theory or design currently exists in the literature. The following overview of the history and current status of stereo and surround sound technology, and particularly of ambisonics, shows that it is particularly appropriate that development of the second-order soundfield microphone should receive more attention at the present time.

1.1.1: The Development of Stereo

Experimental and theoretical studies of directional psychoacoustics were carried out, in particular by Lord Rayleigh (John Strutt), during the latter part of the 19th century. Rayleigh was able to establish a number of results of fundamental importance. Specifically, he determined that at low frequencies the phase difference between the pressure waveforms at the ears is the primary "cue" utilised to localise the sound source, while at higher frequencies, where the head is large enough to result in significant acoustic shadowing, the interaural intensity difference is employed. This so-called "duplex" theory remains the most basic and most important model of directional perception. Rayleigh also noted that additional spatial information may be obtained by movement of the head. Furthermore, he conjectured that the pinnae contribute to localisation by modifying the spectra of sounds impinging on the ear in a manner dependent on the direction of incidence; this is now known to be the case [86] [87] [88] [89].

The earliest demonstrations of a sound reproduction system capable of preserving directional information seem to have been those conducted by Clément Ader in Paris during 1881, in which microphones were positioned along the edge of the stage of the Paris Grand Opera House, and their outputs fed to pairs of telephones held to the ears of listeners [55] [3].

Despite this, it appears that little if any work on the development of stereo technology took place before the 1930's. At that time, it was well known that if the signals obtained from two microphones separated by a distance approximately equal to that between the ears were used to drive a pair of headphones, so that each ear was presented with the output of just one microphone, then an accurate impression of spatial location was obtained. However, during the 1920's, it had become customary to use loudspeakers instead of headphones for the reproduction of sound; this is a fundamentally different situation, since the output of each loudspeaker is received by both ears of the listener. Any loudspeaker-based system must take account of this interaural crosstalk [3] [17] [63].

Perhaps the most important work on stereo recording and reproduction was that conducted by Alan Blumlein at EMI during the 1930's, which forms the basis of stereo systems to the present day. Blumlein believed that the most important directional cue was the low frequency interaural phase difference, and the system which he devised is based upon the observation that, if two loudspeakers are driven by in-phase signals, then at low frequencies the resulting sound pressure waveforms at the ears of a suitably positioned listener differ in phase by an amount that depends upon the amplitudes of the loudspeaker signals; the perceived localisation can be controlled by suitably adjusting these amplitudes. Appropriate signals may be obtained by using pairs of coincident microphones with suitable directional responses. This system is described in Blumlein's frequently cited patent [12], although the earliest publication of the underlying theory appears to have been [17]. While Blumlein's system employed two transmission channels and two loudspeakers, it is clear from [12] that he also recognised the potential desirability of using a greater number of loudspeakers.

Contemporaneously, research into stereophonic sound was also undertaken at Bell Telephone Laboratories. This work differed from Blumlein's in that multi-channel systems were studied as well as two-channel systems, and widely spaced microphones rather than coincident arrangements were employed; also, the Bell Laboratories researchers considered the interaural intensity difference to be the primary localisation cue. In 1933, the Bell Laboratories team together with the Philadelphia Orchestra conducted a public demonstration of stereo reproduction; a performance by the orchestra, located in the auditorium of the American Academy of Music in Philadelphia, was relayed via a three-microphone, three-channel, three-loudspeaker system to an audience in Constitution Hall, Washington [34] [85] [10] [17] [32]. Reportedly, the members of the audience were deeply impressed by the quality of the reproduction obtained [3] [34]; indeed, the Bell Laboratories system produced results which were described as being superior to anything previously heard [17].

It was not possible in the 1930’s to develop commercially viable stereophonic replay equipment for the domestic market, because no suitable two-channel (or multi-channel) medium existed for the distribution of recordings. In fact, it was not until the 1950’s that appropriate distribution media became available; systems utilising two-track reel-to-reel magnetic tape were available by 1956 or 1957 [17] [71], and the stereo phonograph disc was introduced in 1958 [19].

During this period it was thought that addition of the capability to include directional information was the most significant improvement that could be made to the then-current monophonic systems, offering greater benefits than further enhancements in frequency response, distortion characteristics or signal-to-noise ratio [17] [67]. It is interesting to observe that some consideration was already being given to the desirability of extending the reproduced sound stage to surround the listener [91]. However, no technique had been developed to achieve this using a two-channel two-loudspeaker system [17], while domestic and economic considerations, as well as the two-channel nature of the available distribution media, prohibited the use of either more channels or more loudspeakers.

Two-channel two-loudspeaker stereo reproduction has been found to produce results that are subjectively very acceptable, albeit with certain limitations. Convincing inter-speaker, or “phantom”, images can not usually be obtained if the angle subtended by the loudspeakers at the listener’s position exceeds approximately 60°; larger angular separation usually results in a “hole in the middle” effect whereby central images are drawn to one of the loudspeakers, or take on a “recessed” or distant quality. The reproduced sound stage tends to collapse to the nearest loudspeaker if the listener moves very far from a position equidistant from the two loudspeakers. Apparent source locations may move when the listener rotates his head, and phantom sources sometimes appear to be elevated above the loudspeakers.

Various theoretical explanations for these and other aspects of stereo listening have been proposed [17] [41] [44] [49] [63] [64] [75] [85]; there exists some significant degree of disagreement and indeed contradiction between the different theories. This is probably due in part to that fact that all two-channel stereo systems to some extent present the ears with conflicting directional information [41]. Furthermore, the attributes of the reproduced sound are strongly dependent on the microphone arrangement used for recording; since there is no one standard microphone technique, so it is not usually possible to make general statements about reproduction that will hold for all recordings.

1.1.2: Binaural & Transaural Stereo

It has, as previously mentioned, long been known that spatial information is effectively conveyed when the signals obtained from two microphones, separated by approximately the width of the human head, are played through headphones so that the signal from each microphone is communicated to just one ear. This approach to recording and playback is termed "binaural"; the principle is to record the sound pressure at the positions that would have been occupied by the listener's ears had he been physically present in the original sound field, and then reproduce the same pressure variations at his ears on playback. Superior results may be obtained if a "dummy head" is used, with microphones positioned inside the ears, since the effect of the head and outer ears on the incident sound field is then included in the recording. Binaural signals may also be synthesised by convolving source signals with the appropriate "head-related transfer functions" (HRTF's).

Research into the use of binaural and dummy head techniques is ongoing, but despite the extremely impressive results that can be obtained using such systems, they have never achieved widespread popularity. It appears that the use of headphones is not popular with the majority of consumers, and this may be a significant barrier to the general acceptance of binaural recordings [75]. The reproduced sound stage also suffers from certain problems, most particularly front-back ambiguities or transpositions and a tendency for intracranial localisation; these are perhaps at least in part due to mismatches between the dummy head used for recording (or the HRTF's used for synthesis) and the head of the listener. In addition, the sound stage necessarily moves with the listener's head, which is an unnatural situation and means that head movements can not be used to resolve ambiguities [60].

Binaural recordings give poor results when replayed over loudspeakers, because the necessary channel separation cannot be maintained in the presence of interaural crosstalk. So-called "transaural" stereo systems are designed to circumvent this problem by pre-conditioning the loudspeaker feed signals with the inverse of the anticipated crosstalk, so that the original binaural signals are correctly received at the ears [9] [75].

The possibility that systems of this type could provide full periphonic surround sound using just two channels and two loudspeakers makes them potentially extremely attractive. The earliest experiments with transaural stereo did not result in the concept being generally accepted, because of the sensitivity to listener position and room acoustics of the systems used [75]. While later work has allowed these problems to be significantly alleviated, the ultimate practicality of such systems may be regarded as not yet having been fully evaluated, and rather more work seems to be required before they can be considered viable for general use. In addition, it is by no means clear that transaural playback will rectify all of the problems of localisation associated with binaural reproduction. It should however be noted that transaural techniques are not restricted exclusively to the presentation of binaural recordings; a two-loudspeaker transaural system could also, for example, be used to reproduce multi-channel surround sound material via an array of "virtual loudspeakers" [9] [75].

1.1.3: Quadraphonic Systems

By the late 1960's / early 1970's, it was thought that four-loudspeaker pantophonic surround sound systems could be successfully marketed, and several such systems, collectively known as "quadraphonic" (or "quad") systems, were introduced. Commercial pressure resulted in systems being released without sufficient research and development; consequently, they proved to be seriously flawed [15] [21] [30] [32] [41] [46].

Quadraphonic systems were based on the assumption that a separate audio channel should be provided to drive each of the four loudspeakers. The conventional studio practice was to treat the four loudspeakers as four stereo pairs, located in front of, behind, and to either side of the listener, and to assign sound sources to interloudspeaker directions by dividing the signal between the two loudspeakers constituting the appropriate pair. However, while this "pairwise mixing" technique gives satisfactory results with two-channel two-loudspeaker stereo, where the loudspeakers are in front of the listener and subtend an angle of no more than 60°, it is considerably less acceptable with the 90° loudspeaker separation imposed by the use of a square layout. Furthermore, regardless of the loudspeaker separation, pairwise mixing gives poor localisation when the loudspeakers are behind the listener, and fails almost entirely for loudspeaker pairs at the sides. Thus this approach, which was not supported by any psychoacoustic justification, consistently produced inadequate results [21] [30] [31] [41] [45] [46].

It is more difficult to make general statements about the performance of quadraphonic systems in reproducing an original sound field, since there existed no one well-defined standard microphone technique, and a wide range of configurations were used. However, these arrangements were not usually designed in any systematic manner, and so it is not surprising that the results obtained were again often poor [42] [45].

A second fundamental problem with quadraphonic systems was that the signals conveyed to the listener were loudspeaker feed signals. These signals were expected to drive a square loudspeaker layout, since this was considered to be the standard for quadraphonic reproduction. Consequently, if the listener's loudspeakers were not arranged in a square, the directional effect created would not be that which was intended. In many domestic situations, a rectangular layout is more convenient than a square [15] [32] [45].

The third significant problem for quadraphonic systems was that of conveying four independent signals using the available two-channel media. The most common supposed solution to this was so-called "4-2-4 matrixing"; the loudspeaker feed signals were matrixed to form two transmission signals, and it was pretended that the four separate signals could be recovered at the reproduction stage by means of a second matrixing operation. This is obviously impossible, since linear matrixing of two signals can never produce more than two independent signals. Genuinely independent loudspeaker feed signals could therefore not be obtained; high separation could be obtained between some of the loudspeaker feeds, but not between all of them. A further difficulty was presented by the desirability of maintaining backward compatibility with existing reproduction equipment; ideally, the two signals should have been such that an acceptable effect was obtained when they were reproduced via two loudspeakers with no quadraphonic decoding, and also when the two signals were summed for monophonic reproduction. The various systems differed in the extent to which they failed to fulfil these requirements [46].

Signal-actuated parametric (SAP) decoders, also known variously as “logic”, “vario-matrix” or “gain-riding” decoders, were introduced in an attempt to compensate for the poor signal separation resulting from the use of 4-2-4 matrixing. Such decoders employ variable channel gains, which are dynamically adjusted to favour the direction in which the dominant sound source should be localised. This technique can indeed give a sharp directional effect for dominant sound sources, although it can also cause quieter sources to “wander”, but it is only of use when one source is clearly dominant; much of the time, this is not the case. In addition, SAP decoders are tiring to listen to for extended periods, and the resulting “listening fatigue” is not conducive to musical appreciation; such decoders are also incapable of reproducing the subtleties of the reverberant sound field. SAP decoders are therefore not capable of producing satisfactory results [41] [45] [46].

Some quadraphonic systems were designed to avoid the problems inherent in matrixing by utilising instead vinyl discs carrying two conventional audio signals along with two additional signals encoded by frequency modulation of ultrasonic sub-carriers. This technique did allow four genuinely independent loudspeaker feeds to be conveyed, thereby eliminating the problem of poor signal separation, but special equipment was required to extract the additional signals. Furthermore, such discs tended to be noisier and more prone to faulty manufacture than ordinary stereo discs, and the ultrasonic signals were “delicate”, so that they tended to wear out. Nevertheless, these “discreet” quadraphonic systems were sometimes capable of producing impressive directional effects, although they were still subject to the limitations imposed by the fundamental flaws discussed above [21] [46].

Given the number of faults inherent in the various quadraphonic systems, it is perhaps not surprising that statements such as “[quadraphonics] offered various combinations of problems with, occasionally, some interesting effects” [21], “producers can decide for themselves which collection of faults they prefer” [46] and “[quadraphonics] set domestic surround sound back by about 15 years” [59] may be found in the literature. While it would be incorrect to suggest that no quadraphonic material was considered to possess any artistic merit, there was generally widespread dissatisfaction with the poor performance of quadraphonic systems, and although these systems could generate impressive but superficial directional effects, the majority of listeners decided that conventional two-channel stereo usually gave more aesthetically pleasing results. Because of this, and also the incompatibility between the different 4-2-4 matrix systems, none of the quadraphonic systems achieved any lasting popularity or significant commercial success; by the end of the 1970's, they were viewed simply as failed attempts at surround sound [2].

It is notable that there has been no subsequent major attempt by the recording industry to introduce a surround sound system. (The specification for CD allows for four-channel surround sound, but no four-channel CD players have been manufactured and no four-channel discs have been issued [23].)

1.1.4: Surround Sound in the Cinema

By contrast with its failure in the domestic market, multi-channel surround sound has proved rather successful in the cinema.

The first film with a recorded soundtrack was *The Jazz Singer*, released in October 1927. Contrary to the prior expectations of the film industry, *The Jazz Singer* was immensely successful; consequently, by the end of the 1920's almost 8000 theatres in the USA were equipped to display "talkies". In 1940, *Fantasia* became the first film to be released with a multi-channel stereo soundtrack. The so-called "FantaSound" format utilised three channels: left wall / screen / right wall.

Both Warner Brothers and 20th-Century Fox deployed four-channel left screen / centre screen / right screen / surround effects formats in the 1950's. The purpose of the centre channel in such formats is to stabilise and reinforce the centre image; few patrons in a typical cinema auditorium are seated even approximately equidistant between the left and right loudspeakers, and it is undesirable for the central image (usually including all the dialogue) to be "collapsed" to one side or the other, as would otherwise tend to happen. The surround effects channel is not intended to aid the localisation of apparent sound sources, but rather to facilitate effects intended to increase the "excitement" of the presentation; a realistic sound stage is not a primary requirement for an audience already engaged in "suspension of disbelief".

These formats required the soundtrack to be recorded magnetically, instead of employing the standard optical sound-on-film encoding, which could not accommodate four channels. Since not all cinemas were equipped to reproduce such magnetically recorded soundtracks, it was necessary to produce prints carrying optically encoded monophonic sound as well as the multi-channel surround sound prints. The expense of maintaining this dual inventory situation, in addition to the high costs of the required cinema equipment and the surround sound prints themselves, led to the rapid demise of these systems [2].

Cinerama utilised five-channel, six-channel and even seven-channel formats for some 70 mm wide-screen releases, but these failed commercially; the extra appeal to the audience of wide screen and surround sound was not sufficient to justify their expense. Hence, multi-channel surround sound had virtually disappeared from the cinema by the early 1970’s [2].

The successful and permanent re-introduction of surround sound to the cinema was finally accomplished in 1977 with the release of *Star Wars*, which carried a Dolby Stereo soundtrack. Although this was not the first film to utilise Dolby Stereo, it was the massive success of *Star Wars* which encouraged the installation of the decoding equipment in theatres.

Dolby Stereo is similar in concept to the quadraphonic 4-2-4 matrix systems (and in fact is based on one such system). Left, centre, right, and surround effects signals are matrixed to produce just two signals, which can be optically encoded on 35 mm film; a Dolby Stereo matrix decoder processes these to generate approximations to the original four signals. The encoded signals produce acceptable results when replayed over a two-channel system with no matrix decoding, so that backward compatibility with two-channel stereo replay systems is maintained. Although such matrix systems are unsatisfactory when used for audio-only reproduction, Dolby Stereo proved to be quite acceptable in the cinema where accurate localisation and a realistic sound stage are not of primary importance; indeed, the processing applied to the surround signal in the Dolby Stereo system is specifically intended to result in diffuse and ambiguous sound, the source of which cannot be localised [5] [84].

In the 1980’s, an enhancement to Dolby Stereo was produced in the form of Dolby Pro Logic decoders. These are SAP decoders, which allow more dramatically impressive effects to be obtained from the encoded Dolby Stereo signals. Once again, a technique unsuitable for audio-only applications is quite satisfactory for cinematic use.

The Dolby Digital system was introduced in 1992. This does not use matrixing, being instead a discrete "5.1" channel system providing five independent full-bandwidth digital channels for left, right, centre, left surround and right surround signals, along with a limited-bandwidth low frequency effects (LFE) channel. To maintain backward compatibility and avoid any need for dual inventory, all Dolby Digital film prints also carry a two-channel analogue Dolby Stereo soundtrack; Dolby's AC-3 lossy compression algorithm is employed to reduce the data rate so that the six digital signals can be optically encoded within the available space on the print.

In 1999, Dolby deployed an extended version of this system, Dolby Digital Surround-EX, which includes a centre surround (i.e., rear centre) signal matrixed into the left and right surround channels (it is interesting to note that *Star Wars Episode 1: The Phantom Menace* was the first film to utilise this format).

Two other discrete multi-channel digital surround sound formats have become reasonably widely supported during the last decade; DTS (Digital Theatre Systems) and SDDS (Sony Dynamic Digital Sound). However, neither of these has achieved either the level of consumer recognition or the degree of market penetration enjoyed by the Dolby systems.

1.1.5: "Home Cinema"

Private screening facilities have been available to the sufficiently wealthy for some considerable time; towards the end of the 1980's, however, the "home theatre" or "home cinema" became a commercially viable mass-market concept [2].

In 1978 the stereo-capable video cassette recorder was introduced, and shortly afterwards Dolby Surround decoders were also made available (Dolby Surround is simply Dolby Stereo renamed for the domestic market, probably to avoid the two-channel two-loudspeaker connotations of the word "stereo"). This was significant because many of the films released on video had originally been produced using Dolby Stereo, and the same soundtracks were used for the video releases; something made possible by the backward compatibility of Dolby Stereo with two-channel two-loudspeaker reproduction. Assisted considerably by consumer recognition of the Dolby brand name, the release of Dolby Surround decoders was a significant factor in the eventual success of home cinema.

However, the first such decoders were significantly inferior to the theatrical units, tending to give rather poor results. This was rectified in 1987, when Dolby Pro Logic decoders for the domestic market were first produced. In addition, stereo television broadcasting commenced during the mid-1980's, and a number of programmes utilised Dolby Surround encoded soundtracks. These developments helped secure the success of the home cinema concept, and also ensured that surround sound was an essential part of that concept [2].

Hence, multi-channel surround sound reproduction systems finally became established as commercially successful domestic products, albeit only in support of visual media, and not (as yet) for audio-only applications. Contemporary consumer Dolby Pro Logic systems typically use five full-range loudspeakers, one for each of the left, centre and right channels and two, positioned rear left and rear right, for the surround signal. Some decoders may also derive a low frequency feed to drive an optional subwoofer for enhanced bass.

During the mid-1980's there were suggestions that Dolby Surround, as the de facto surround sound standard, could be used for material other than film and television soundtracks. This prospect was viewed with considerable alarm by the small group of extant surround sound audio enthusiasts, since Dolby Surround is not at all suited to audio-only use [84]. To the author's knowledge, such proposals were never implemented.

1.1.6: DVD

Recently, the DVD (Digital Versatile Disc) has been introduced as a distribution medium for the domestic entertainment market. Two forms of DVD are of interest here; DVD-Video (DVD-V), and DVD-Audio (DVD-A). At the time of writing, DVD-V is clearly a successful consumer product; by contrast, it appears to the author that DVD-A remains almost unknown outside of the professional audio community.

DVD-V is obviously an ideal format for digital home cinema systems. DVD-A is intended as a "super CD" audio-only format, providing higher sampling frequencies and increased word lengths as well as multi-channel sound. DVD-V of course also supports multi-channel surround sound, since this is an essential part of the home cinema concept.

Because so much of the storage capacity of a DVD-V disc is required to store the visual material, the multi-channel audio content has to be encoded using lossy compression. Since Dolby Digital is the primary format used for digital multi-channel audio with lossy compression in the cinema, it is no surprise that this is the most common format utilised on DVD-V discs (although other formats, such as DTS, are also supported by the DVD-V standard). It is equally unsurprising that 5.1 channel audio is therefore now considered to be the standard for domestic multi-channel sound, and that the preferred loudspeaker layout for the reproduction of such material has become the conventional domestic configuration [16] [22].

DVD-A discs are, of course, not required to store video material. However, some of the more demanding options supported by the DVD-A specification, such as six-channel audio sampled at 96 kHz with a word length of 24 bits and a playing time of at least 74 minutes, involve so much data that the capacity of the disc is exceeded with PCM encoding. Since mandatory use of lossy compression is not acceptable on a high quality audio disc, a lossless compression scheme known as "Meridian Lossless Packing" (MLP) (sometimes also referred to as "packed PCM") is part of the DVD-A specification [24] [35].

It is interesting to observe that there exist audio professionals, recording engineers and "audiophile" listeners who believe that the capability to deliver multi-channel surround sound is a much more important feature of DVD than the provision for increased word length or sample rate, although it should be noted that this view is by no means universally held [16] [24] .

1.1.7: Ambisonics

Ambisonic surround sound systems were initially developed in the early 1970's. They were, however, significantly different from any of the flawed quadraphonic systems, having exactly the kind of rational psychoacoustic basis that was so conspicuously lacking from those systems.

Ambisonic systems are based on a mathematical model of directional psychoacoustics developed by Michael Gerzon. Gerzon believed that the design of surround sound systems had previously not been a systematic procedure, and that there was little possibility of achieving optimal results while this remained the case. The model he developed, which subsumed many previous theories of localisation, was specifically intended to describe directional psychoacoustics in a mathematical form that could conveniently be used in calculations relating to surround sound system design [41] [44] [46] [49]. Ambisonic systems do not exploit all known localisation mechanisms, but they do take into account more directional cues than other systems, resulting in benefits such as reduced listening fatigue, good interloudspeaker imaging with reduced susceptibility to the "detent effect" (the tendency for apparent source locations to be "pulled" towards the closest loudspeaker) and improved image stability.

There exists a hierarchy of ambisonic systems, described as first-order, second-order, and so on. The precise meaning of "first-order" in this context is discussed in Chapter 4. A system of higher order is capable of creating superior directional effects and providing more precise localisation of sound sources than a system of lower order; however, systems of higher order require a greater number of signals [5] [6] [31] [40].

Periphony was part of the initial conception of ambisonic systems. The periphonic capabilities of ambisonics are seen by some as a significant advantage; it has been claimed that height information "adds almost as much to the realism of a surround system as rear speakers do" [22], or even that "periphonic reproduction is ... at least as great an advance on pantophonic as the latter is on stereo, or stereo on mono" [31]. Unsurprisingly, periphonic systems require more signals than pantophonic systems.

Ambisonics employs a signal set, known as "B-format", which is based on the principle of encoding direction, without reference to the loudspeaker layout used for reproduction. The B-format signal set is described in Chapter 4. As a result of this, ambisonic systems are adaptable to multiple loudspeaker layouts; an ambisonic decoder derives appropriate loudspeaker feed signals from the transmitted B-format signals. To obtain good performance from ambisonic systems, the number of loudspeakers should exceed the number of B-format signals used; a further increase in the number of loudspeakers will usually give improved results [15] [45] [90]. Periphonic systems, not surprisingly, require more loudspeakers than pantophonic systems; higher-order systems also require a greater number of loudspeakers.

During the initial development of ambisonics, concentration was very much focussed on first-order pantophonic systems. Although some publications did include material relating to periphonic and / or higher-order systems [31] [40], only first-order pantophonic systems were considered to be commercially viable at the time. Such systems require three signals and a minimum of four loudspeakers. A two-channel matrix system known as UHJ was also developed to allow the use of two-channel distribution media and provide vital two-channel stereo compatibility; although this inevitably gave poorer results than three-channel B-format, it was still superior to the quadraphonic 4-2-4 matrix systems. The UHJ signal format is discussed further in Appendix 2.

With a two-channel distribution format and a four-loudspeaker reproduction system, ambisonics should have been viable in a market that had already been prepared to accept inferior quadraphonic systems. However, it did not achieve any significant commercial success. This failure has been attributed to “money [and] bureaucratic bungling of the kind only the English can manage” [22] (though not, it should be made clear, on the part of the inventors); a rather more detailed account may be found in [21].

Despite this, researchers and audio professionals have never entirely abandoned ambisonics, and another opportunity to effectively market the system has long been hoped for. The arrival of DVD may provide this opportunity.

The DVD-V specification does not include explicit support for ambisonics. Neither, despite a certain amount of lobbying [1], does the DVD-A specification per se. However, the MLP system itself does have the capabilitry to carry ambisonic material [54]. This may be useful in the future, once DVD-A players with the capability to pass the encoded MLP bit stream to an external decoder become available.

A method has also been devised for conveying ambisonic recordings on DVD-V or DVD-A discs without the need for a decoder or any explicit support for the format. This may be done by pre-decoding the ambisonic recording for the typical 5.1 channel loudspeaker configuration; more details of this so-called “G-format” technique may be found in Appendix 2.

1.1.8: The Soundfield Microphone

Much of the early work on ambisonics was concerned with the capture of live performances (although the importance of modern multi-track studio techniques was also acknowledged, and appropriate ambisonic studio equipment designed [21] [45]). This required the use of a suitable microphone technique. The necessary signals can in principle be obtained by employing a coincident array of microphones, somewhat similar to the coincident pairs pioneered by Blumlein, but consisting of four rather than two microphone capsules.

In practise, it is not possible to take four separate microphones and place them sufficiently close together. The first-order soundfield microphone was therefore developed for this purpose. It consists of a tetrahedral array of first-order microphones together with signal processing apparatus which produces a good approximation to the signals that would have been obtained from the ideal coincident array [28] [31] [42] [53]. This is described in detail in Chapter 5.

Unlike ambisonic reproduction systems, the first-order soundfield microphone has proved to be a successful product, because it is also useful for making conventional two-channel stereo recordings [20] [94].

A first-order soundfield microphone can only produce first-order B-format signals. Live capture of second-order B-format signals required for second-order ambisonics reproduction will require a second-order soundfield microphone. Just as the first-order soundfield microphone has found application in the field of conventional two-channel stereo recording, so too may the second-order microphone. In particular, the narrower pickup patterns associated with second-order microphones could be beneficial when recordings are to be made in particularly reverberant environments, or when the microphones must be positioned far from the sound sources [51].

Soundfield microphones can also be used to advantage when making recordings intended for reproduction via non-ambisonic multi-channel surround sound systems. The more highly directional responses attainable using a second-order soundfield microphone could be of benefit when making recordings for five-channel surround sound systems [37].

1.2: Overview of Thesis Content

In Chapter 2 of this thesis, some essential mathematical background information is presented. In particular, two families of special functions, the spherical harmonics and the spherical Bessel functions, are described; these are of great importance in the analysis of soundfield microphones and ambisonic systems.

In Chapter 3, the theory of first-order and second-order pressure gradient microphones is developed. The author believes that the existing treatments are somewhat inadequate; starting with the concept of the directional derivative as the appropriate mathematical abstraction for the operation of a pressure gradient microphone, the presentation here is systematic and unified. Regardless of any such benefit, the presentation of a general theory of second-order pressure gradient microphones is in any case unique, since only certain special cases have received attention in the literature.

Aspects of ambisonic systems are considered in Chapter 4. In particular, the B-format signal set is described and the relationship betwen its usual defintion in terms of spherical harmonics and a more fundamental basis in terms of pressure gradients is explained. Some aspects of ambisonic reproduction are examined; this analysis extends that which has previously been published. Some largely unrecognised assumptions underlying the definitions of B-format and ambisonic systems are examined, and it is demonstrated that these assumptions may in fact result in the standard definitions being suboptimal.

In Chapter 5, the theory of the first-order soundfield is presented. It is not easy to obtain a comprehensive understanding of this theory from the existing literature; the presentation is intended to “fill in the gaps”.

The theory and design of the second-order soundfield microphone is developed in Chapter 6. This material is entirely original.

Finally, in Chapter 7 an alternative signal set formulation for second-order systems is proposed; this is free from the assumptions referred to above, and may therefore be capable of offering superior performance.

2: Mathematical Preliminaries

The coordinate conventions employed in this thesis are those which are customary in the literature relating to ambisonics and soundfield microphones. Where cartesian coordinates are used, the coordinate frame is oriented such that, from the point of view of an observer at the origin, the positive x axis points forwards, the positive y axis points to the left, and the positive z axis points upwards. Unit vectors in the x, y and z directions are denoted by $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$. Spherical polar coordinates are specified in terms of azimuth angle q, measured anti-clockwise from the positive x axis, and elevation angle f, measured from the x-y plane with positive values corresponding to points in the upper half-space. The coordinates of a point in one system may be found in terms of its coordinates in the other using the relationships

$$r = \sqrt{x^2 + y^2 + z^2} \qquad (2.1a)$$

$$q = \arctan(y/x) \qquad (2.1b)$$

$$f = \arctan\left(z \middle/ \sqrt{x^2 + y^2}\right) \qquad (2.1c)$$

and

$$x = r\cos(q)\cos(f) \qquad (2.2a)$$

$$y = r\sin(q)\cos(f) \qquad (2.2b)$$

$$z = r\sin(f) \qquad (2.2c)$$

It should be noted that other spherical polar coordinate conventions are also in common use, both in mathematical work generally, and specifically in the audio engineering literature; unfortunately, these often assign different meanings to the same symbols q and f.

The following definitions of the forward and inverse Fourier transforms are used:

$$\hat{F}\{f(t)\} = F(w) = \int_{-\infty}^{\infty} f(t) e^{-jwt} dt \qquad (2.3a)$$

$$\hat{F}^{-1}\{F(w)\} = f(t) = \frac{1}{2p} \int_{-\infty}^{\infty} F(w) e^{jwt} dw \qquad (2.3b)$$

2.1: Spherical Harmonics & The Laplace Series

The polar response pattern of a microphone is a function of direction; it may therefore be thought of as a function defined on the surface of a sphere. There exists a series representation for such functions, somewhat analogous to the Fourier series expansion of a function defined along a line; the series is known as the Laplace series, and the basis functions are termed spherical harmonics. Microphone polar patterns are often discussed in terms of their Laplace series representation, especially in the context of soundfield microphones and ambisonic systems.

A direction in three-dimensional space may be specified by the azimuth and elevation angles q and f or by the three direction cosines

$$\hat{x} = \cos(q)\cos(f) \tag{2.4a}$$

$$\hat{y} = \sin(q)\cos(f) \tag{2.4b}$$

$$\hat{z} = \sin(f) \tag{2.4c}$$

An *n*th-order spherical harmonic is, by definition, any homogeneous polynomial function of degree *n* in the direction cosines that is orthogonal with respect to integration over the surface of the unit sphere to all spherical harmonics of lower order [40] [56]. Any function on the sphere may be uniquely expressed as the sum of a zeroth-order spherical harmonic, a first-order spherical harmonic, a second-order spherical harmonic, ..., an *m*th-order spherical harmonic, ... (This statement is not strictly true, but conditions for the existence of the series [56] are satisfied by most functions likely to be encountered in practice and, in particular, by all functions considered in this thesis.)

It may be shown that, for any given order *n*, it is possible to define exactly $2n+1$ linearly independent spherical harmonics; all other spherical harmonics of the same order may then be expressed as a sum of these [40] [56]. Since

$$\sum_{m=0}^{n}(2m+1) = (n+1)^2 \tag{2.5}$$

so the cumulative total number of linearly independent spherical harmonics of all orders up to and including ***n*** is $(n+1)^2$ [40] [56].

In principle, for any order ***n*** an infinite number of different sets of $2n+1$ linearly independent spherical harmonics may be defined. There exists however a preferred or canonical set [56], defined in terms of Legendre's polynomials and Legendre's associated functions:

$$P_n(\sin(f))$$
$$\cos(mq)P_n^m(\sin(f)) \quad 1 \le m \le n$$
$$\sin(mq)P_n^m(\sin(f)) \quad 1 \le m \le n$$

where $P_n(m)$ is the Legendre's polynomial (or Legendre's function) of degree ***n*** and $P_n^m(m)$ is the associated Legendre's function of the first kind of order ***m*** and degree ***n*** for $-1 \le m \le 1$. The Legendre's polynomial of degree ***n*** can be found using Rodrigues' formula [56]

$$P_n(m) = \frac{1}{2^n n!} \frac{d^n}{dm^n} \left\{ \left(m^2 - 1 \right)^n \right\} \tag{2.6}$$

The associated Legendre's functions are defined by [56]

$$P_n^m(m) = \left(1 - m^2 \right)^{m/2} \frac{d^m P_n(m)}{dm^m} \tag{2.7}$$

where $\left(1 - m^2\right)^{m/2}$ takes its positive value when ***m*** is odd. Hence, when $m = \sin(f)$,

$$P_n^m(\sin(f)) = \cos^m(f) \frac{d^m P_n(\sin(f))}{d(\sin(f))^m} \tag{2.8}$$

The material presented in this thesis involves spherical harmonics of order up to and including six; the Legendre's polynomials and Legendre's associated functions are given for $n \le 6$ in tables 2.1 and 2.2. The zeroth-order, first-order and second-order spherical harmonics are of particular importance, and are given in table 2.3; spherical harmonics of higher order can easily be obtained using the definitions above and the information presented in tables 1 and 2.

Spherical harmonics of the form $P_n(\sin(\phi))$ are termed zonal harmonics; the remaining spherical harmonics are known as tesseral harmonics, except when $m = n$, when they are called sectorial harmonics. There are therefore two sectorial harmonics of every order $n \geq 1$; these take the forms $\cos(n\theta)\cos^n(\phi)$ and $\sin(n\theta)\cos^n(\phi)$.

In some branches of physics the so-called Condon-Shortley phase convention, in which an extra factor of $(-1)^m$ is included in the definitions either of the spherical harmonics themselves or of the Legendre's associated functions, is standard [4]. This convention is not usually followed in the ambisonics literature, and is not adopted in this thesis. (Note however that it is implicitly assumed in [56].)

Given a function on the sphere $F(\theta,\phi)$, the Laplace series expansion is

$$F(\theta,\phi) = \sum_{n=0}^{\infty} A_n P_n(\sin(\phi)) + \sum_{n=1}^{\infty}\sum_{m=1}^{n}\left(A_{n,m}\cos(m\theta) + B_{n,m}\sin(m\theta)\right)P_n^m(\sin(\phi)) \qquad (2.9)$$

where the coefficients are given by

$$A_n = \frac{2n+1}{4\pi}\int_0^{2\pi}\int_{-\pi/2}^{\pi/2} P_n(\sin(\phi'))F(\theta',\phi')\cos(\phi')\,d\phi'\,d\theta' \qquad (2.10a)$$

$$A_{n,m} = \frac{2n+1}{2\pi}\frac{(n-m)!}{(n+m)!}\int_0^{2\pi}\int_{-\pi/2}^{\pi/2}\cos(m\theta')P_n^m(\sin(\phi'))F(\theta',\phi')\cos(\phi')\,d\phi'\,d\theta' \qquad (2.10b)$$

$$B_{n,m} = \frac{2n+1}{2\pi}\frac{(n-m)!}{(n+m)!}\int_0^{2\pi}\int_{-\pi/2}^{\pi/2}\sin(m\theta')P_n^m(\sin(\phi'))F(\theta',\phi')\cos(\phi')\,d\phi'\,d\theta' \qquad (2.10c)$$

The scaling factors in these expressions arise because, while the spherical harmonics are orthogonal, they are not orthonormal.

The terminology "*n*th-order spherical harmonic" is used in this thesis, in accordance with the existing literature concerning ambisonics; however, such functions are formally known as "ordinary spherical surface harmonics of degree *n*" (distinguishing them from others which are also called spherical harmonics) [56].

n	$P_n(m)$
0	1
1	m
2	$\frac{1}{2}(3m^2 - 1)$
3	$\frac{1}{2}(5m^3 - 3m)$
4	$\frac{1}{8}(35m^4 - 30m^2 + 3)$
5	$\frac{1}{8}(63m^5 - 70m^3 + 15m)$
6	$\frac{1}{16}(231m^6 - 315m^4 + 105m^2 - 5)$

Table 2.1: Legendre's Polynomials

m	$P_1^m(m)$	$P_2^m(m)$	$P_3^m(m)$	$P_4^m(m)$
1	$(1-m^2)^{1/2}$	$3m(1-m^2)^{1/2}$	$\frac{1}{2}(1-m^2)^{1/2}(15m^2-3)$	$\frac{5}{2}m(1-m^2)^{1/2}(7m^2-3)$
2		$3(1-m^2)$	$15m(1-m^2)$	$\frac{15}{2}(1-m^2)(7m^2-1)$
3			$15(1-m^2)^{3/2}$	$105m(1-m^2)^{3/2}$
4				$105(1-m^2)^2$

m	$P_5^m(m)$	$P_6^m(m)$
1	$\frac{15}{8}(1-m^2)^{1/2}(21m^4-14m^2+1)$	$\frac{21}{8}m(1-m^2)^{1/2}(33m^4-30m^2+5)$
2	$\frac{105}{2}m(1-m^2)(3m^2-1)$	$\frac{105}{8}(1-m^2)(33m^4-18m^2+1)$
3	$\frac{105}{2}(1-m^2)^{3/2}(9m^2-1)$	$\frac{105}{2}m(1-m^2)^{3/2}(33m^2-9)$
4	$945m(1-m^2)^2$	$\frac{945}{2}(1-m^2)^2(11m^2-1)$
5	$945(1-m^2)^{5/2}$	$10395m(1-m^2)^{5/2}$
6		$10395(1-m^2)^3$

Table 2.2: Legendre's Associated Functions of the First Kind for $-1 \le m \le 1$

n, *m*	Laplace Series Coefficient	Spherical Harmonic
0, 0	A_0	1
1, 0	A_1	$\sin(f)$
1, 1	$A_{1,1}$ $B_{1,1}$	$\cos(q)\cos(f)$ $\sin(q)\cos(f)$
2, 0	A_2	$\tfrac{1}{2}\left(3\sin^2(f)-1\right)\equiv\tfrac{1}{4}\left(1-3\cos(2f)\right)$
2, 1	$A_{2,1}$ $B_{2,1}$	$3\cos(q)\cos(f)\sin(f)\equiv\tfrac{3}{2}\cos(q)\sin(2f)$ $3\sin(q)\cos(f)\sin(f)\equiv\tfrac{3}{2}\sin(q)\sin(2f)$
2, 2	$A_{2,2}$ $B_{2,2}$	$3\cos(2q)\cos^2(f)$ $3\sin(2q)\cos^2(f)$

Table 2.3: Spherical Harmonics of Order $n \le 2$

2.2: Spherical Bessel Functions

The spherical Bessel functions may be defined in terms of the Bessel functions of order $n+\tfrac{1}{2}$, where *n* is an integer. Specifically, the spherical Bessel function of the first kind of order *n* is defined as [92]

$$j_n(a)=\sqrt{\frac{p}{2a}}\,J_{n+\frac{1}{2}}(a) \tag{2.11}$$

It follows from the properties of Bessel functions of order $n+\tfrac{1}{2}$ that each spherical Bessel function may be written as the weighted sum of a finite number of terms of the form $\sin(a)/a^q$ and $\cos(a)/a^q$, where the integer $q \le n+1$ [92].

Spherical Bessel functions satisfy the three-term recurrence relationship

$$j_{n-1}(a)+j_{n+1}(a)=\frac{2n+1}{a}\,j_n(a) \tag{2.12}$$

and their derivatives may be expressed in the form

$$\frac{d}{da}\{j_n(a)\}=\frac{1}{2n+1}\left[nj_{n-1}(a)-(n+1)j_{n+1}(a)\right] \tag{2.13}$$

The spherical Bessel functions of orders up to and including seven, required in this thesis, are shown in table 2.4.

It may be shown that

$$j_n(0)=\begin{cases}1 & n=0\\ 0 & \text{otherwise}\end{cases} \tag{2.14}$$

A result of particular importance in the theory of soundfield microphones is [42] [68] [92]

$$e^{ja\cos(b)}=\sum_{n=0}^{\infty} j^n(2n+1)P_n\big(\cos(b)\big)j_n(a) \tag{2.15}$$

n	$j_n(a)$
0	$\frac{\sin(a)}{a}$
1	$\frac{\sin(a)}{a^2}-\frac{\cos(a)}{a}$
2	$3\frac{\sin(a)}{a^3}-3\frac{\cos(a)}{a^2}-\frac{\sin(a)}{a}$
3	$15\frac{\sin(a)}{a^4}-15\frac{\cos(a)}{a^3}-6\frac{\sin(a)}{a^2}+\frac{\cos(a)}{a}$
4	$105\frac{\sin(a)}{a^5}-105\frac{\cos(a)}{a^4}-45\frac{\sin(a)}{a^3}+10\frac{\cos(a)}{a^2}+\frac{\sin(a)}{a}$
5	$945\frac{\sin(a)}{a^6}-945\frac{\cos(a)}{a^5}-420\frac{\sin(a)}{a^4}+105\frac{\cos(a)}{a^3}+15\frac{\sin(a)}{a^2}-\frac{\cos(a)}{a}$
6	$10395\frac{\sin(a)}{a^7}-10395\frac{\cos(a)}{a^6}-4275\frac{\sin(a)}{a^5}+1260\frac{\cos(a)}{a^4}+210\frac{\sin(a)}{a^3}$ $-21\frac{\cos(a)}{a^2}-\frac{\sin(a)}{a}$
7	$135135\frac{\sin(a)}{a^8}-135135\frac{\cos(a)}{a^7}-62370\frac{\sin(a)}{a^6}+17325\frac{\cos(a)}{a^5}+3150\frac{\sin(a)}{a^4}$ $-378\frac{\cos(a)}{a^3}-28\frac{\sin(a)}{a^2}+\frac{\cos(a)}{a}$

Table 2.4: Spherical Bessel Functions

2.2.1: Inverse Fourier Transforms of Spherical Bessel Functions

In the analysis of soundfield microphones to follow, frequency responses will arise which are conveniently expressed in terms of spherical Bessel functions of frequency. To establish the corresponding impulse responses, it is of course necessary to know the inverse Fourier transforms.

Since the spherical Bessel function of order zero is the well-known “sinc” function, so its inverse Fourier transform is a standard result:

$$\hat{F}^{-1}\{j_0(Kw)\} = \frac{1}{2K} r_K(t) \tag{2.16}$$

where

$$r_K(t) = \begin{cases} 1 & -K < t < K \\ 0 & \text{otherwise} \end{cases} \tag{2.17}$$

It follows as a special case of equation (2.13) that

$$\begin{aligned} j_1(Kw) &= -\frac{d}{d(Kw)}\{j_0(Kw)\} \\ &= -\frac{1}{K}\frac{d}{dw}\{j_0(Kw)\} \end{aligned} \tag{2.18}$$

and therefore

$$\begin{aligned} \hat{F}^{-1}\{j_1(Kw)\} &= -\frac{1}{K}\hat{F}^{-1}\left\{\frac{d}{dw}\{j_0(Kw)\}\right\} \\ &= -\frac{1}{K}(-jt)\hat{F}^{-1}\{j_0(Kw)\} \\ &= j\frac{1}{2K^2} t r_K(t) \end{aligned} \tag{2.19}$$

Now, by rearranging equation (2.13), we obtain

$$\begin{aligned} j_m(Kw) &= \frac{m-1}{m} j_{m-2}(Kw) - \frac{2m-1}{m}\frac{d}{d(Kw)}\{j_{m-1}(Kw)\} \\ &= \frac{m-1}{m} j_{m-2}(Kw) - \frac{2m-1}{m}\frac{1}{K}\frac{d}{dw}\{j_{m-1}(Kw)\} \end{aligned} \tag{2.20}$$

and taking inverse Fourier transforms of both sides gives

$$\hat{F}^{-1}\{j_m(Kw)\} = \frac{m-1}{m}\hat{F}^{-1}\{j_{m-2}(Kw)\} + j\frac{2m-1}{mK} t\hat{F}^{-1}\{j_{m-1}(Kw)\} \tag{2.21}$$

Thus, given the inverse Fourier transforms of $j_0(Kw)$ and $j_1(Kw)$, the inverse transforms of $j_2(Kw)$ and spherical Bessel functions of higher order may be found by repeated application of equation (2.21). The inverse Fourier transforms of spherical Bessel functions of order up to and including seven are shown in table 2.5.

n	$\hat{F}^{-1}\{j_n(Kw)\}$
0	$\frac{1}{2K} r_K(t)$
1	$j\frac{1}{2K^2} t r_K(t)$
2	$-\frac{1}{4K^3}\left[3t^2 - K^2\right] r_K(t)$
3	$-j\frac{1}{4K^4}\left[5t^2 - 3K^2\right] t r_K(t)$
4	$\frac{1}{16K^5}\left[35t^4 - 30K^2t^2 + 3K^4\right] r_K(t)$
5	$j\frac{1}{16K^6}\left[63t^4 - 70K^2t^2 + 15K^4\right] t r_K(t)$
6	$-\frac{1}{96K^7}\left[693t^6 - 945K^2t^4 + 315K^4t^2 - 15K^6\right] r_K(t)$
7	$-j\frac{1}{96K^8}\left[1287t^6 - 2079K^2t^4 + 945K^4t^2 - 105K^6\right] t r_K(t)$

Table 2.5: Inverse Fourier Transforms of Spherical Bessel Functions