Radio Communication Handbook

ELEVENTH EDITION

Editors
Mike Dennison, G3XDV
John Fielding, ZS5JF

Radio Society of Great Britain

Published by the Radio Society of Great Britain, 3 Abbey Court, Fraser Road, Priory Business Park, Bedford. MK44 3WH. Tel 01234 832700. Web www.rsgb.org

First published 2011.

Cover design: Kim Meyern

Production: Mark Allgar, M1MPA

Design and layout: Mike Dennison, G3XDV, Emdee Publishing

Printed in Great Britain by Latimer Trend of Plymouth

Companion CD printed by DBMasters of Faversham, England (www.dbmasters.co.uk)

ISBN 9781-9050-8674-0

Contents

Note: Many chapters have references to the RSGB Bulletin, Radio Communication or RadCom. These are historic names of the RSGB members' monthly journal. The magazines are available on a series of CD-ROMs from: RSGB, 3 Abbey Court, Fraser Road, Priory Business Park, Bedford. MK44 3WH (www.rsgb.org)

6 HF Receivers

Roger Wilkins, G8NHG

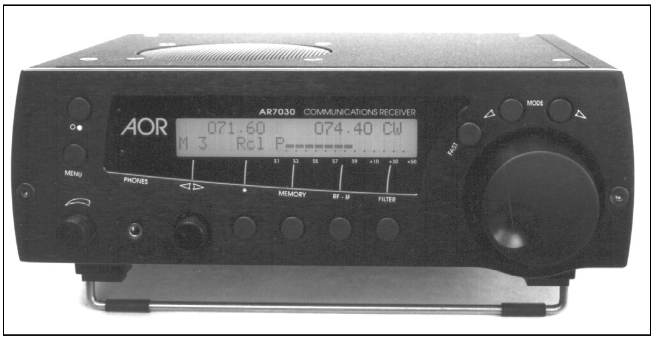


Fig 6.1: The AOR 7030 is a sophisticated receiver covering the frequency range 0-32MHz

Amateur HF operation, whether for two-way contacts or for listening to amateur transmissions, imposes stringent requirements on the receiver. The need is for a receiver that enables an experienced operator to find and hold extremely weak signals on frequency bands often crowded with much stronger signals from local stations or from the high-power broadcast stations using adjacent bands. The wanted signals may be fading repeatedly to below the external noise level, which limits the maximum usable sensitivity of HF receivers, and which will be much higher than in the VHF and UHF spectrum.

Although the receivers now used by most amateurs form part of complex, factory-built HF transceivers, the operator should understand the design parameters that determine how well or how badly they will perform in practice, and appreciate which design features contribute to basic performance as HF communications receivers, as opposed to those which may make them more user-friendly but which do not directly affect the reception of weak signals. This also applies to dedicated receivers that are factory-built, such as the one shown in Fig 6.1.

Ideally, an HF receiver should be able to provide good intelligibility from signals which may easily differ in voltage delivered from the antenna by up to 10,000 times and occasionally by up to one million times (120dB) - from less than 1µV from a weak signal to nearly 1V from a near-neighbour. **Table 1** shows the relationship between the various ways of measuring the input signals; pd (potential difference) and dBm (input power) are most commonly used.

To tune and listen to SSB or to a stable CW transmission while using a narrow-band filter, the receiver needs to have a frequency stability of within a few hertz over periods of 15 minutes or so, representing a stability of better than one part in a million. It should be capable of being tuned with great precision, either continuously or in increments of at most a few hertz.

A top-quality receiver may be required to receive transmissions on all frequencies from 1.8MHz to 30MHz (or even 50MHz) to provide 'general coverage' or only on the bands allotted to amateurs. Such a receiver may be suitable for a number of different modes of transmission - SSB, CW, AM, NBFM, data (RTTY/packet) etc - with each mode imposing different requirements in selectivity, stability and demodulation (decoding). Such a receiver would inevitably be complex and costly to buy or build.

On the other hand, a more specialised receiver covering only a limited number of bands and modes such as CW-only or CW/SSB-only, and depending for performance rather more on the skill of the operator, can be relatively simple to build at low cost.

As with other branches of electronics, the practical implementation of high-performance communications receivers has undergone a number of radical changes since their initial development in the mid-1930s, some resulting from the improved stability needed for SSB reception and others aimed at reducing costs by substituting electronic techniques in place of mechanical precision.

However, it needs to be emphasised that, in most cases, progress in one direction has tended to result in the introduction of new problems or the enhancement of others: "What we call progress is the exchange of one nuisance for another nuisance" (Havelock Ellis) or "Change is certain; progress is not" (A J P Taylor). As late as 1981, an Australian amateur was moved to write: "Solid-state technology affords commercial manufacturers cheap, large-scale production but for amateur radio receivers and transceivers of practical simplicity, valves remain incomparably superior for one-off, home-built projects."

The availability of linear integrated circuits capable of forming the heart of communications receivers combined with the increasing scarcity and hence cost of special valve types has tended to reverse this statement. It is still possible to build reasonably effective HF receivers, particularly those for limited frequency coverage, on the kitchen table with the minimum of test equipment.

for a 50 ohms power matched system	emf	pd	dBm
	1 V	0.5 V	7.0
	100 mV	50 mV	-13.0
	10 mV	5 mV	-33.0
	1 mV	0.5 mV	-53.0
possible range of signal levels	100 uV	50 uV	-80.5
	10 uV	5 uV	-93.0
	1 uV	0.5 uV	-113.0
typical receiver sensitivity			
thermal noise floor (3KHz bandwidth)	0.1 uV	50 nV	-133.0

Table 6.1: The relationship between emf, pd and dBm

Furthermore, since many newcomers will eventually acquire a factory-built transceiver but require a low-cost, stand-alone HF receiver in the interim period, the need can be met either by building a relatively simple receiver, or by acquiring, and if necessary modifying, one of the older valve-type receivers that were built in very large numbers for military communications during the second world war, or those marketed for amateur operation in the years before the virtually universal adoption of the transceiver.

Even where an amateur has no intention of building or servicing his or her own receiver, it is important that he or she should have a good understanding of the basic principles and limitations that govern the performance of all HF communications receivers.

BASIC REQUIREMENTS

The main requirements for a good HF receiver are:

- Sufficiently high sensitivity, coupled with a wide dynamic range and good linearity to allow it to cope with both the very weak and very strong signals that will appear together at the input; it should be able to do this with the minimum impairment of the signal-to-noise ratio by receiver noise, cross-modulation, blocking, intermodulation, reciprocal mixing, hum etc.
- Good selectivity to allow the selection of the required signal from among other (possibly much stronger) signals on adjacent or near-adjacent frequencies. The selectivity characteristics should 'match' the mode of transmission, so that interference susceptibility and noise bandwidth should be as close as possible to the intelligence bandwidth of the signal.
- Maximum freedom from spurious responses - that is to say signals which appear to the user to be transmitting on specific frequencies when in fact this is not the case. Such spurious responses include those arising from image responses, breakthrough of signals and harmonics of the receiver's internal oscillators.
- A high order of stability, in particular the absence of short-term frequency drift or jumping.
- Good read-out and calibration of the frequency to which the set is tuned, coupled with the ability to reset the receiver accurately and quickly to a given frequency or station.
- Means of receiving SSB and CW, normally requiring a stable beat frequency oscillator preferably in conjunction with product detection.
- Sufficient amplification to allow the reception of signals of under 1μV input; this implies a minimum voltage gain of about one million times (120dB), preferably with effective automatic gain control (AGC) to hold the audio output steady over a very wide range of input signals.
- Sturdy construction with good-quality components and with consideration given to problems of access for servicing when the inevitable occasional fault occurs.

A number of other refinements are also desirable: for example it is normal practice to provide a headphone socket on all communications receivers; it is useful to have ready provision for receiver 'muting' by an externally applied voltage to allow voice-operated, push-to-talk or CW break-in operation; an S-meter to provide immediate indication of relative signal strengths; a power take-off socket to facilitate the use of accessories; an IF signal take-off socket to allow use of external special demodulators for NBFM, FSK, DSBSC, data etc.

In recent years, significant progress has continued to be made in meeting these requirements - although we are still some way short of being able to provide them over the entire signal range of 120dB at the ideal few hertz stability. The introduction of more and more semiconductor devices into receivers has brought a number of very useful advantages, but has also paradoxically made it more difficult to achieve the highly desirable wide dynamic range. Professional users now require frequency read-out and long-term stability of an extremely high order (better than 1Hz stability is needed for some applications) and this has led to the use of frequency synthesised local oscillators and digital read-out systems; although these are effective for the purposes which led to their adoption, they are not necessarily the correct approach for amateur receivers since, unless very great care is taken, a complex frequency synthesiser not only adds significantly to the cost but may actually result in a degradation of other even more desirable characteristics.

So long as continuous tuning systems with calibrated dials were used, the mechanical aspects of a receiver remained very important; it is perhaps no accident that one of the outstanding early receivers (HRO) was largely designed by someone whose early training was that of a mechanical engineer.

It should be recognised that receivers which fall far short of ideal performance by modern standards may nevertheless still provide entirely usable results, and can often be modified to take advantage of recent techniques. Despite all the progress made in recent decades, receiver designs dating from the 'thirties and early 'forties are still capable of being put to good use, provided that the original electrical and mechanical design was sound. Similarly, the constructor may find that a simple, straightforward and low-cost receiver can give good results even when its specification is well below that now possible. It is ironical that almost all the design trends of the past 30 years have, until quite recently, impaired rather than improved the performance of receivers in the presence of strong signals!

BASIC TYPES OF RECEIVERS

Amateur HF receivers fall into one of two main categories:

(a) 'straight' regenerative and direct-conversion receivers in which the incoming signal is converted directly into audio by means of a demodulator working at the signal frequency;
(b) single- and multiple-conversion superhet receivers in which the incoming signal is first converted to one or more intermediate frequencies before being demodulated. Each type of receiver has basic advantages and disadvantages.

Regenerative Detector ('Straight' or TRF) Receivers

At one time valve receivers based on a regenerative (reaction) detector, plus one or more stages of AF amplification (ie 0-V-1, 0-V-2 etc), and sometimes one or more stages of RF amplification at signal frequency (1-V-1 etc) were widely used by amateurs. High gain can be achieved in a correctly adjusted regenerative detector when set to a degree of positive feedback just beyond that at which oscillation begins; this makes a regenerative receiver capable of receiving weak CW and SSB signals. However, this form of detector is non-linear and cannot cope well in situations where the weak signal is at all close to a strong signal; it is also inefficient as an AM detector since the gain is much reduced when the positive feedback (regeneration) is reduced below the oscillation threshold. Since the detector is non-linear, it is usually impossible to provide adequate selectivity by means of audio filters.

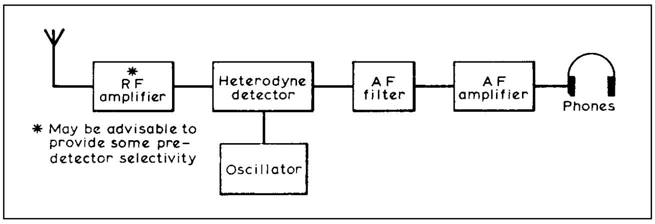


Fig 6.2: Outline of a simple direct-conversion receiver in which high selectivity can be achieved by means of audio filters

Simple Direct-conversion Receivers

A modified form of 'straight' receiver which can provide good results, even under modern conditions, becomes possible by using a linear detector which is in effect simply a frequency converter, in conjunction with a stable local oscillator set to the signal frequency (or spaced only the audio beat away from it). Provided that this stage has good linearity in respect of the signal path, it becomes possible to provide almost any desired degree of selectivity by means of audio filters (Fig 6.2).

This form of receiver (sometimes termed a homodyne) has a long history but only in the last few decades of the 20th century did it become widely used for amateur operation since it is more suited (in its simplest form) to CW and SSB reception than AM. The direct-conversion receiver may be likened to a superhet with an IF of 0kHz or alternatively to a straight receiver with a linear rather than a regenerative detector.

In a superhet receiver the incoming signal is mixed with a local oscillator signal and the intermediate frequency represents the difference between the two frequencies; thus as the two signals approach one another the IF becomes lower and lower. If this process is continued until the oscillator is at the same frequency as the incoming signal, then the output will be at audio (baseband) frequency; in effect one is using a frequency changer or translator to demodulate the signal. Because high gain cannot be achieved in a linear detector, it is necessary to provide very high AF amplification. Direct-conversion receivers can be designed to receive weak signals with good selectivity but in this form do not provide true single-sideband reception (see later); another problem often found in practice is that very strong broadcast signals (eg on 7MHz) drive the detector into non-linearity and are then demodulated directly and not affected by any setting of the local oscillator.

A crystal-controlled converter can be used in front of a direct-conversion receiver, so forming a superhet with variable IF only. Alternatively a frequency converter with a variably tuned local oscillator providing output at a fixed IF may be used in front of a direct-conversion receiver (regenerative or linear demodulator) fixed tuned to the IF output. Such a receiver is sometimes referred to as a supergainer receiver.

Two-phase and 'Third-method' Direct-Conversion Receivers

An inherent disadvantage of the simple direct-conversion receiver is that it responds equally to signals on both sides of its local oscillator frequency, and cannot reject what is termed the audio image no matter how good the audio filter characteristics; this is a serious disadvantage since it means that the selectivity of the receiver can only be made half as good as the theoretically ideal bandwidth. This problem can be overcome, though at the cost of additional complexity, by phasing techniques similar to those associated with SSB generation. Two main approaches are possible: see Fig 6.3.

Fig 6.3(a) shows the use of broad-band AF 45 degree phase-shift networks in an 'outphasing' system, and with care can result in the reduction of one sideband to the extent of 30-40dB. Another possibility is the polyphase SSB demodulator which does not require such critical component values as conventional SSB phase-shift networks.

Fig 6.3(b) shows the 'third method' (sometimes called the Weaver or Barber system) which requires the use of additional balanced mixers working at AF but eliminates the need for accurate AF phase-shift networks. The 'third method' system, particularly in its AC-coupled form [1] provides the basis for high-performance receivers at relatively low cost, although suitable designs for amateur operation are rare. Two-phase direct-conversion receivers based on two diode-ring mixers in quadrature (90° phase difference) are capable of the high performance of a good superhet.

HF Superhet Receivers

The vast majority of receivers are based on the superhet principle. By changing the incoming signals to a fixed frequency (which may be lower or higher than the incoming signals) it becomes possible to build a high-gain amplifier of controlled selectivity to a degree which would not be possible over a wide spread of signal frequencies. The main practical disadvantage with this system is that the frequency conversion process involves unwanted products which give rise to spurious responses, and much of the design process has to be concentrated on minimising the extent of these spurious responses in practical situations.

A single-conversion superhet is a receiver in which the incoming signal is converted to its intermediate frequency, amplified

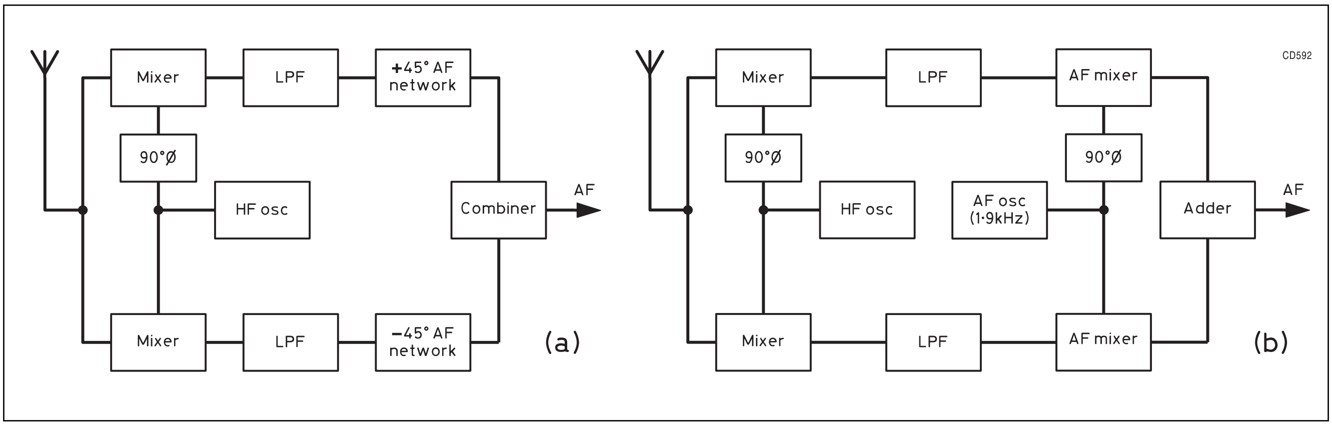
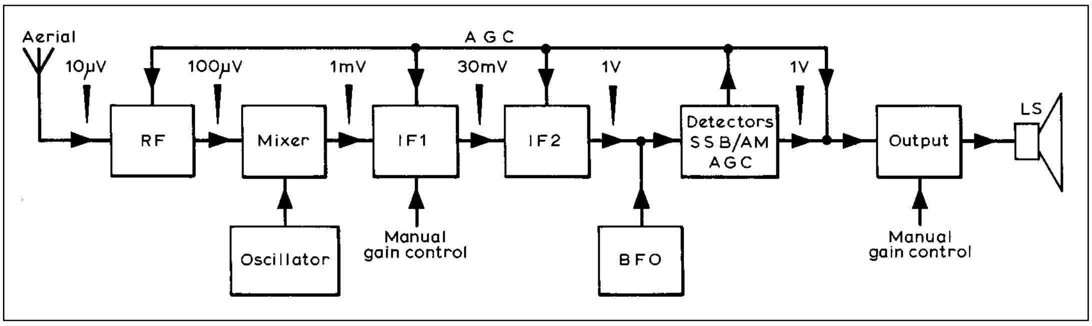


Fig 6.3: Block outline of two-phase ('autophasing') form of direct-conversion receiver. (b) Block outline of 'third method' (Weaver or Barber) SSB direct-conversion receiver

Fig 6.4: Block outline of representative single-conversion



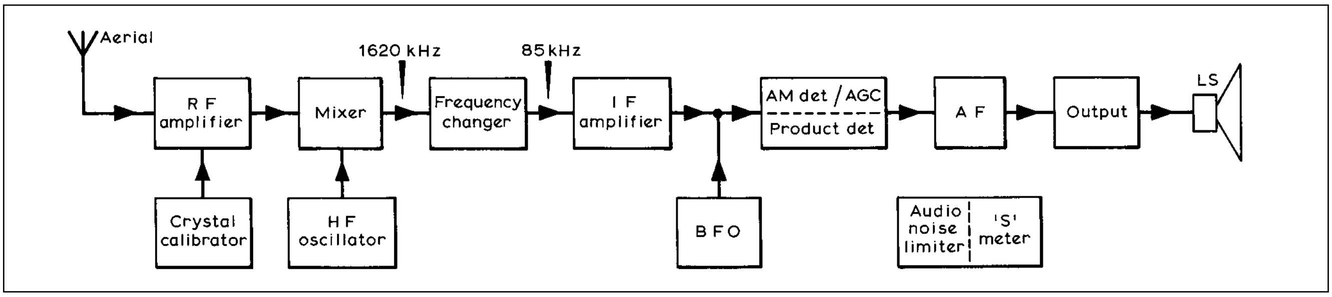


Fig 6.5: Block outline of double-conversion communications receiver with both IFs fixed

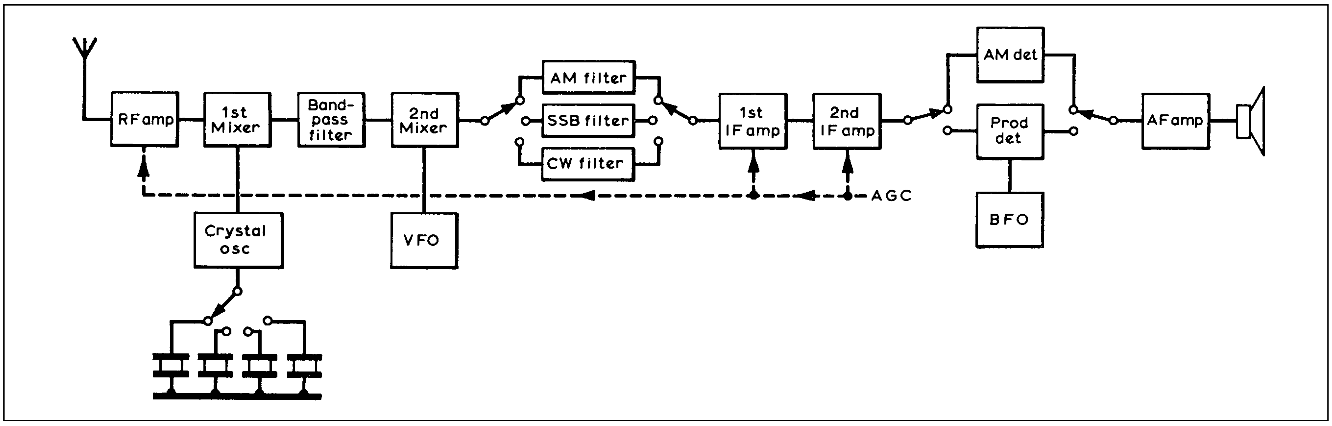


Fig 6.6: Block diagram of a double-conversion receiver with crystal-controlled first oscillator - typical of many late 20th-century designs

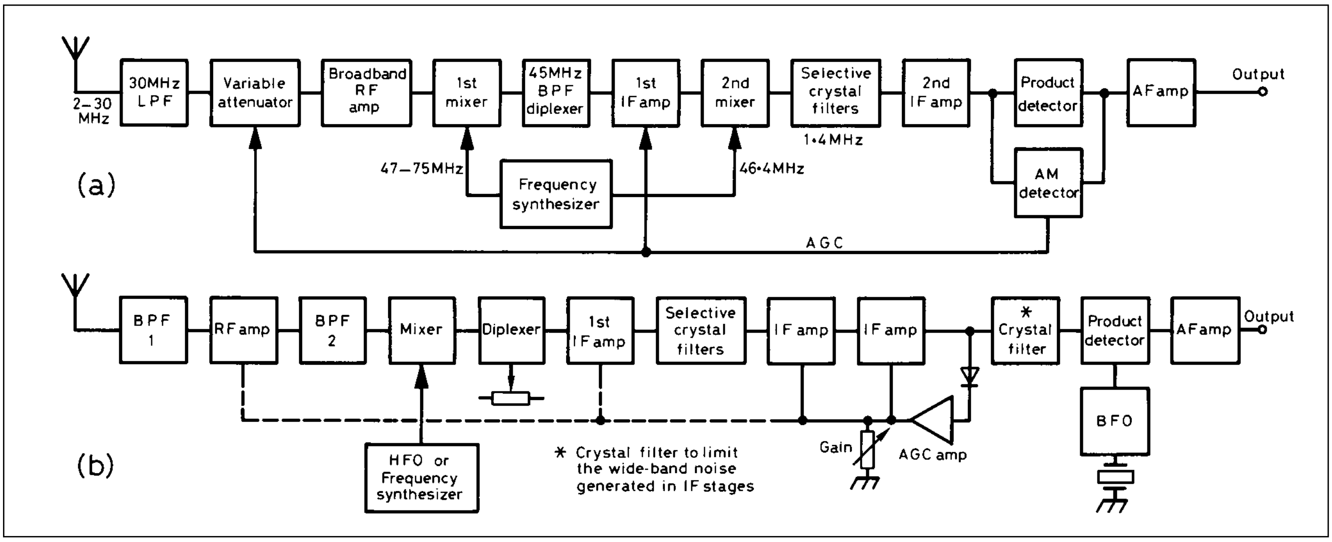


Fig 6.7: Representative architectures of modern communications receiver designs. (a) General-coverage double-conversion superhet with up-conversion to 45MHz first IF and 1.4MHz second IF. (b) Single-conversion superhet, typically for amateur bands only, with an IF in the region of 9 or 10MHz

and then demodulated at this second frequency. Virtually all domestic AM broadcast receivers use this principle, with an IF of about 455-470kHz, and a similar arrangement but with refinements was found in many communications receivers. However, for reasons that will be made clear later in this chapter, some receivers convert the incoming signal successively to several different frequencies; these may all be fixed IFs: for example the first IF might be 9MHz and the second 455kHz and possibly a third at 35kHz. Or the first IF may consist of a whole spectrum of frequencies so that the first IF is variable when tuning a given band, with a subsequent second conversion to a fixed IF. There are in fact many receivers using double or even triple conversion, and a few with even more conversions, though unless care is taken each conversion makes the receiver susceptible to more spurious responses. The block diagram of a typical single-conversion receiver is shown in **Fig 6.4**. **Fig 6.5** illustrates a double-conversion receiver with fixed IFs, while **Fig 6.6** is representative of a receiver using a variable first IF in conjunction with a crystal-controlled first local oscillator (HFO).

Many modern factory-built receivers up-convert the signal frequency to a first IF at VHF as this makes it more convenient to use a frequency synthesiser as the first HF oscillator: **Fig 6.7(a)**. As the degree of selectivity provided in a receiver increases, it reaches the stage where the receiver becomes a single-sideband receiver, although this does not mean that only SSB signals can be received.

In fact the first application of this principle was the single-signal receiver for CW reception where the selectivity is sufficient to reduce the strength of the audio image (resulting from beating the IF signal with the BFO) to an insignificant value, thus virtually at one stroke halving the apparent number of CW stations operating on the band (previously each CW signal was heard on each side of the zero beat). Similarly, double-sideband AM signals can be received on a set having a carefully controlled pass-band as though they were SSB, with the possibility of receiving either sideband should there be interference on the other. This degree of selectivity can be achieved with good IF filters or alternatively the demodulator can itself be designed to reject one or other of the sidebands, by using phasing techniques similar to those sometimes used to generate SSB signals and for two-phase direct-conversion receivers. But most receivers rely on the use of crystal or mechanical filters to provide the necessary degree of sideband

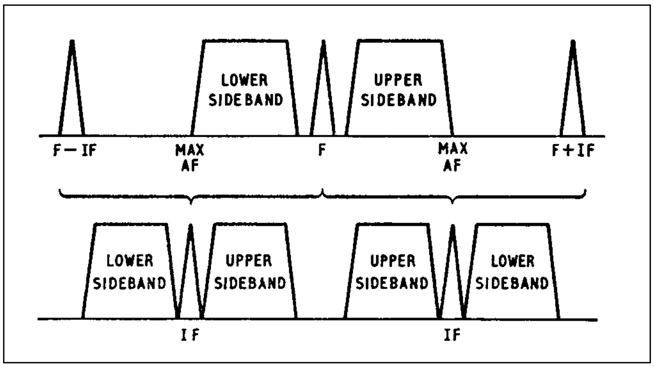


Fig 6.8: A local oscillator frequency lower than the signal frequency (ie f - IF) keeps the upper and lower sidebands of the intermediate frequency signals in their original positions. However, when the local oscillator is placed higher in frequency than the signal frequency (f + IF), the positions of the sidebands are transposed. By incorporating two oscillators, one above and the other below the input signal, sideband selection is facilitated (this is generally carried out at the final IF by switching the BFO or carrier insertion oscillator)

selectivity, and then use heterodyne oscillators placed either side of the nominal IF to select upper or lower sidebands.

It is important to note that whenever frequency conversion is accomplished by beating with the incoming signal an oscillator lower in frequency than the signal frequency, the sidebands retain their original position relative to the carrier frequency, but when conversion is by means of an oscillator placed higher in frequency than the carrier, the sidebands are inverted. That is to say, an upper sideband becomes a lower sideband and vice versa (see **Fig 6.8**).

Software Defined Radio

The latest developments in professional and amateur radio communications technology are in the field of ‘Software Defined Radio’. The increasing power of modern computers and digital signal processing (DSP) are bringing about unprecedented advances in radio technology and performance which looks set to soon overtake the best available from analogue designs.

Digital signal processing has been a useful technique for some years. Initially the speed of available analogue-to-digital converters (ADC) and processors limited the applicability of DSP to audio frequencies. They were nevertheless extremely useful and provided excellent audio filters, with advanced features such as noise blanking and automatic heterodyne suppression.

As devices improved at an exponential rate, it became feasible to implement the final stages of a receiver from low intermediate frequencies onwards digitally. The term Software Defined Radio is intended to convey the concept that the functionality of the radio is essentially defined by the software running on the computer which implements a significant part of the receiver's stages. This has many advantages, including the ease with which many aspects of the receiver's ‘design’ can altered merely by updating software.

A number of commercial designs are now available for the amateur market and have become extremely popular, such as the Flex Radio SDR1000. In many cases, these designs use a Direct Digital Synthesis (DDS) local oscillator and a mixer direct to baseband (zero IF), creating separate 90° phased I and Q signals. The computing power is provided by the operator's desktop PC, which accepts the I and Q signals from the analogue front end and demodulates them into single sideband audio. Advanced on-screen user interfaces provide the operator with numerous features such as panoramic display of a large segment of band, and incredible control over the filtering and characteristics of the receiver.Yet, the ultimate goal of SDR must be considered to be a direct digitisation of the HF Spectrum, covering 0 - 30MHz direct from the antenna, with the entire receiver implemented digitally by software. The only analogue stage in such a system would be a low-pass filter to prevent deterioration of ADC performance due to interfering signals higher than 30MHz. Until recently, the dynamic range available from ADC devices and the computing power requirements were not able to meet the demands of a true all-digital SDR.

However, the ever marching progress in semiconductor techniques and manufacture have made available high performance ADCs which have now made the all-digital HF receiver a reality. Several all-digital receivers are now available which boast high performance over the entire HF range. Whilst it may be argued that the highest performance in terms of dynamic range, sensitivity and immunity from cross-modulation is still the domain of analogue receivers, or at least analogue front ends, there is no doubt that continued advances will soon allow all-digital receivers to overtake their analogue counterparts.

For more information on SDR, please see the later chapter in this Handbook.

DESIGN TRENDS

After the 'straight' receiver, because of its relatively poor performance and lack of selectivity on AM phone signals, had fallen into disfavour in the mid-1930s, came the era of the superhet communications receiver. Most early models were single conversion designs based on an IF of 455-470kHz, with two or three IF stages, a multi-electrode triode-hexode or pentagrid mixer, sometimes but not always with a separate oscillator valve. This approach made at least one RF amplifying stage essential in order to raise the level of the incoming signal before it was applied to the relatively noisy mixer; two stages were to be preferred since this meant they could be operated in less critical conditions and provided the additional pre-mixer RF selectivity needed to reduce 'image' response on 14MHz and above. Usually a band-switched LC HF oscillator was gang-tuned so as to track with two or three signal-frequency tuned circuits, calling for fairly critical and expensive tuning and alignment systems. These receivers were often designed basically to provide full coverage on the HF band (and often also the MF band), sometimes with a second tuning control to provide electrical bandspread on amateur bands, or with provision (as on the HRO) optionally to limit coverage to amateur bands only. Selectivity depended on the use of good-quality IF transformers (sometimes with a tertiary tuned circuit) in conjunction with a single-crystal IF filter which could easily be adjusted for varying degrees of selectivity and which included a phasing control for nulling out interfering carriers.

In later years, to overcome the problem of image response with only one RF stage, there was a trend towards double- or triple-conversion receivers with a first IF of 1.6MHz or above, a second IF about 470kHz and (sometimes) a third IF about 50kHz.

With a final IF of 50kHz it was possible to provide good single-signal selectivity without the use of a crystal filter.

The need for higher stability than is usually possible with a band-switched HF oscillator and the attraction of a similar degree of band-spreading on all bands has led to the widespread adoption of an alternative form of multi-conversion superhet; in effect this provides a series of integral crystal-controlled converters in front of a superhet receiver (single or double conversion) covering only a single frequency range (for example 5000-5500kHz) This arrangement provides a fixed tuning span (in this example 500kHz) for each crystal in the HF oscillator. Since a separate crystal is needed for each band segment, most receivers of this type are designed for amateur bands only (though often with provision for the reception of a standard frequency transmission, for example on 10MHz); more recently some designs have eliminated the need for separate crystals by means of frequency synthesis, and in such cases it is economically possible to provide general coverage. The selectivity in these receivers is usually determined by a band-pass crystal filter, mechanical filter or multi-pole ceramic filter, a separate filter is being used for SSB, CW and AM reception (although for economic reasons sets may be fitted with only one filter, usually intended for SSB reception). In this system the basic 'superhet' section forms in effect a variable IF amplifier. Examples of this tunable IF architecture are the SS-R1 receiver [2] made by Squires Sanders [3], and the G2DAF design published in RadCom [4] and built by many amateurs, both in the 1960s.

In practice the variable IF type of receiver provides significantly enhanced stability and lower tuning rates on the higher frequency bands, compared with receivers using fixed IF, though it is considerably more difficult to prevent breakthrough of strong signals within the variable IF range, and to avoid altogether the appearance of 'birdies' from internal oscillators. With careful design a high standard of performance can be achieved; the use of multiple conversion (with the selective filter further from the antenna input stage) makes the system less suitable for semiconductors than for valves, particularly where broad-band circuits are employed in the front-end and in the variable IF stage.

There is now a trend back to the use of fixed IF receivers, either with single conversion or occasionally with double conversion (provided that in this case an effective roofing filter is used at the first IF). A roofing filter is a selective filter intended to reduce the number of strong signals passing down an IF chain without necessarily being of such high grade or as narrow bandwidth as the main selective filter.

To overcome the problem of image reception a much higher first IF is used; for amateur band receivers this is often 9MHz since effective SSB and CW filters at this frequency are available. This reduces (though does not eliminate) the need for pre-mixer selectivity; while the use of low-noise mixers makes it possible to reduce or eliminate RF amplification. To overcome frequency stability problems inherent in a single-conversion approach, it is possible to obtain better stability with FET oscillators than was usually possible with valves; another approach is to use mixer-VFO systems (essentially a simple form of frequency synthesis) and such systems can provide identical tuning rates on all bands, though care has to be taken to reduce to a minimum spurious injection frequencies resulting from the mixing process.

To achieve the maximum possible dynamic range, particular attention has to be given to the mixer stage, and it is an advantage to make this a balanced, or double-balanced (see later) arrangement using either double-triodes, Schottky (hot-carrier) diodes or FETs (particularly power FETs).

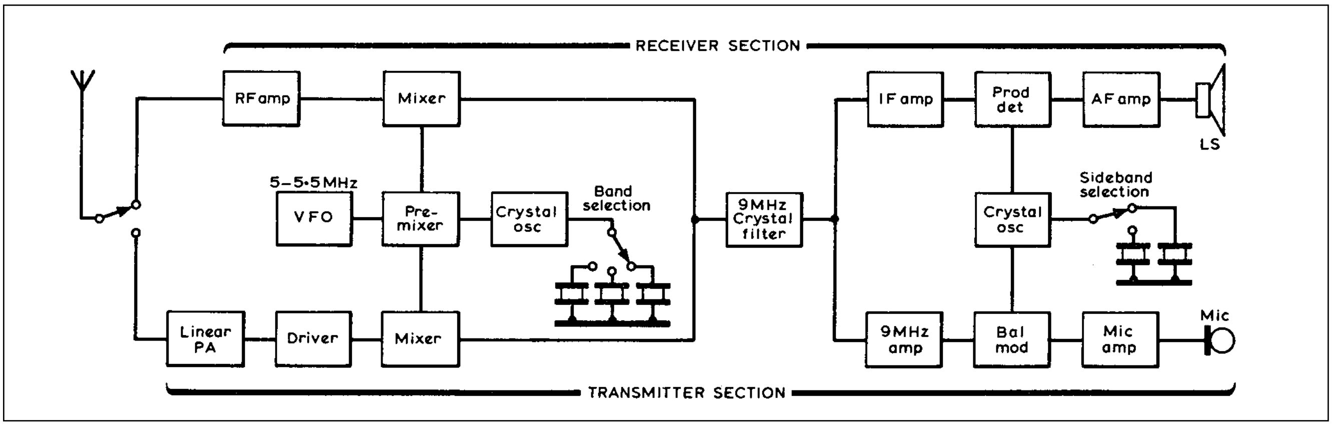


Fig 6.9: Block diagram of a typical modern SSB transceiver in which the receiver is a single-conversion superhet with 9MHz IF in conjunction with the pre-mixer form of partial frequency synthesis

A further significant reduction of spurious responses may prove possible by abandoning the superhet in favour of high-performance direct-conversion receivers (such as the Weaver or 'third-method' SSB direct-conversion arrangement); however, such designs are still only at an early stage of development.

Most modern receivers are built in the form of compact transceivers functioning both as receiver and transmitter, and with some stages common to both functions (**Fig 6.9**). Modern transceivers use semiconductor devices throughout. Dual-gate FET devices are generally found in the signal path of the receiver. Most transceivers have a common SSB filter for receive and transmit; this may be a mechanical or crystal filter at about 455kHz but current models more often use crystal filters at about 3180, 5200, 9000kHz or 10.7MHz, since the use of a higher frequency reduces the total number of frequency conversions necessary.

One of the fundamental benefits of a transceiver is that it provides common tuning of the receiver and transmitter so that both are always 'netted' to the same frequency. It remains, however, an operational advantage to be able to tune the receiver a few kilohertz around the transmit frequency and vice versa, and provision for this incremental tuning is often incorporated; alternatively many transceivers offer two oscillators so that the transmit and receive frequencies may be separated when required.

The most critical aspect of modern receivers is the signal-handling capabilities of the early (front-end) stages. Various circuit techniques are available to enhance such characteristics: for example the use of balanced (push-pull) rather than single-ended signal frequency amplifiers; the use of balanced or double-balanced mixer stages; the provision of manual or AGC-actuated antenna-input attenuators; and careful attention to the question of gain distribution.

An important advantage of modern techniques such as linear integrated circuits and wide-band fixed-tuned filters rather than tuneable resonant circuits is that they make it possible to build satisfactory receivers without the time-consuming and constructional complexity formerly associated with high-performance receivers. Nevertheless a multiband receiver must still be regarded as a project requiring considerable skill and patience.

The widespread adoption of frequency synthesisers as the local HF oscillator has led to a basic change in the design of most factory-built receivers, although low-cost synthesisers may not be the best approach for home-construction. Such synthesisers cannot readily (except under microprocessor control) be 'ganged' to band-switched signal-frequency tuned circuits; additionally, mechanically-ganged variable tuning of band-switched signal-frequency and local oscillator tuned circuits as found in older communications receivers would today be a relatively high-cost technique.

These considerations have led to widespread adoption in factory-built receivers and transceivers of 'single-span' up-conversion multiple-conversion superhets with a first IF in the VHF range, up to about 90MHz, followed by further conversions to lower IFs at which the main selectivity filter(s) are located.

In such designs, pre-selection before the first mixer or preamplifier (often arranged to be optionally switched out of circuit) may simply take the form of a low-pass filter (cut off at 30MHz) or a single wide-band filter covering the entire HF band. Higher-performance receivers usually fit a series of sub-octave band-pass filters, with electronic switching (preferably with PIN diodes).

With fixed filtering, even of the sub-octave type, very strong HF broadcast transmissions will be present at the mixer(s) and throughout the 'front-end' up to the main selectivity filter(s). To enable weak signals to be received free of intermodulation products, this places stringent requirements on the linearity of the front-end. The use of relatively noisy low-cost PLL frequency synthesisers also raises the problem of 'reciprocal mixing' (see later). For home-construction of high-performance receivers, the earlier design approaches are still attractive, including the 'old-fashioned' concept of achieving good pre-mixer selectivity with high-Q tuned circuits using variable capacitors rather than electronic tuning diodes. Diode switching rather than mechanical switching can also significantly degrade the intermodulation performance of receivers. Further, it should be noted that reed relay switching can often introduce sufficient series resistance to seriously degrade the Q of tuned circuits unless reeds especially selected for their RF properties are used.

DIGITAL TECHNIQUES

The availability of general-purpose, low-cost digital integrated circuit devices made a significant impact on the design of communication receivers although, until the later introduction of digital signal processing, their application was primarily for operator convenience and their use for stable, low-cost frequency synthesisers rather than their use in the signal path.

By incorporating a digital frequency counter or by operation directly from a frequency synthesiser, it is now normal practice to display the frequency to which the receiver (or transceiver) is tuned directly on matrices of light-emitting diodes or liquid crystal displays. This requires that the display is offset by the IF from the actual output of the frequency synthesiser or free-running local oscillator. Such displays have virtually replaced the use of calibrated tuning dials.

Frequency synthesisers are commonly 'tuned' by a rotary shaft-encoded switch which can have the 'feel' of mechanical capacitor tuning of a VFO, but this may be supplemented by push-buttons which enable the wanted frequency to be punched

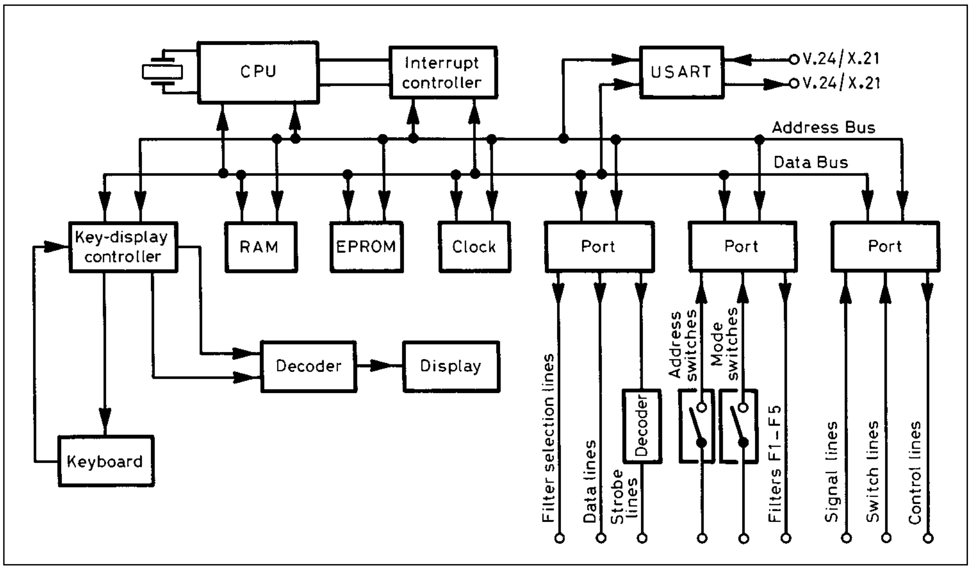


Fig 6.10: Architecture of the elaborate internal computer system found in a modern professional fully-synthesised receiver (DJ2LR)

10 Low Frequencies: Below 1MHz

Jim Moritz, M0BMU

The 136kHz band, 135.7kHz - 137.8kHz was introduced in January 1998 and is unique in being in the LF frequency range (Low Frequency, defined as 30kHz - 300kHz).

In February 2007, the licensing authority, Ofcom. began inviting applications for a Notice of Variation for UK amateurs to operate for experimental and research purposes in the range 501.0kHz - 504.0kHz [1] in the MF range (the Medium Frequency range, defined as 300kHz - 3MHz, and also including the 160m band). The issued NoVs have been renewed periodically by Ofcom; current NoVs are valid until February 2012.

Both 136kHz and "500kHz" bands have unique characteristics and are different to all the higher frequency amateur bands. Propagation on 136kHz and 500kHz is very different to the HF bands (for more on this, see the chapter on propagation).

Due to the narrow bandwidth available at 136kHz (a total of only 2.1kHz), the low radiated power level permitted (1W ERP) and the high noise levels present on this band, several specialised techniques [2, 3] have been developed for 136kHz DX operation, alongside familiar CW for shorter-distance contacts. Some LF DX modes are described in the Digital Communications chapter.

The 500kHz band has a somewhat lower noise level, and the radiated power limit has recently been increased to 10W ERP, but offers its own challenges, particularly the very deep fading that occurs at intermediate and long distances. The majority of 500kHz operation has also been in CW, although digital modes have seen increasing use. In the future, it can be expected that both bands will see the development of new operating modes and techniques to achieve communication under the often very marginal conditions.

RECEIVERS FOR 136kHz AND 500kHz

The majority of stations currently use commercially available receivers. Many amateur HF receivers and transceivers include general coverage that extends to 500kHz and 136kHz, and in many cases these can be successfully pressed into service; however, unlike HF reception, where reasonable results are often achieved simply by connecting a 'random' wire antenna to the receiver input, successful reception on these bands is a bit more difficult, for a number of reasons. First, there is usually a very large mismatch between the impedance of a wire antenna at these frequencies and the typically 50-ohm receiver input impedance, which leads to a large reduction in signal level at the receiver input. This is often exacerbated by degraded receiver sensitivity at 500kHz and below, particularly in amateur-type equipment. Secondly, amateurs with their relatively tiny radiated signals share the spectrum with vastly more powerful broadcast and utility signals; unless effective filtering is provided, this results in severe problems with overloading at the receiver front end. Fortunately, very satisfactory results can usually be achieved by using quite simple antenna matching, preamplifier, and preselector arrangements, as will be seen later in this section.

LF and MF Receiver Requirements

Requirements for LF receivers depend on the type of operation that is envisaged. Adequate sensitivity is obviously required; the internal receiver noise level should be well below the natural band noise at all times. As a guide, similar figures to those for HF receivers (a few tenths of a microvolt in a CW bandwidth) will suffice. If a large, transmitting-type antenna is used, the signal level will be high enough to allow a receiver with considerably poorer sensitivity to be used. Small loop and whip receiving antennas usually require a dedicated low noise preamplifier (see section on receiving antennas and noise reduction).

A widely used operating mode is CW. Because of the narrow bandwidth available (136kHz is only 2.1kHz wide, and 500kHz is 3kHz wide) a CW filter is almost essential. A 500Hz bandwidth is adequate, but 250Hz or narrower bandwidths can be used to advantage. On 136kHz, there are several strong utility signals just outside (and sometimes inside) the amateur band (**Fig 10.1**), and so good filter shape factor is important since the utility signals can be 60dB or more above the level of readable amateur signals, with a frequency separation of less than 1kHz. For specialised extremely narrow-bandwidth modes such as QRSS (see the chapters on Morse and Digital Communications), selectivity is also provided at audio frequencies using DSP techniques in a personal computer, but good basic receiver selectivity is still desirable to prevent strong out-of-band signals entering the audio stages.

The spectrum around 500kHz was once heavily used for maritime communications, but now non-amateur signals close to the 501kHz - 504kHz range are rare (Fig 10.1), making selectivity less critical.

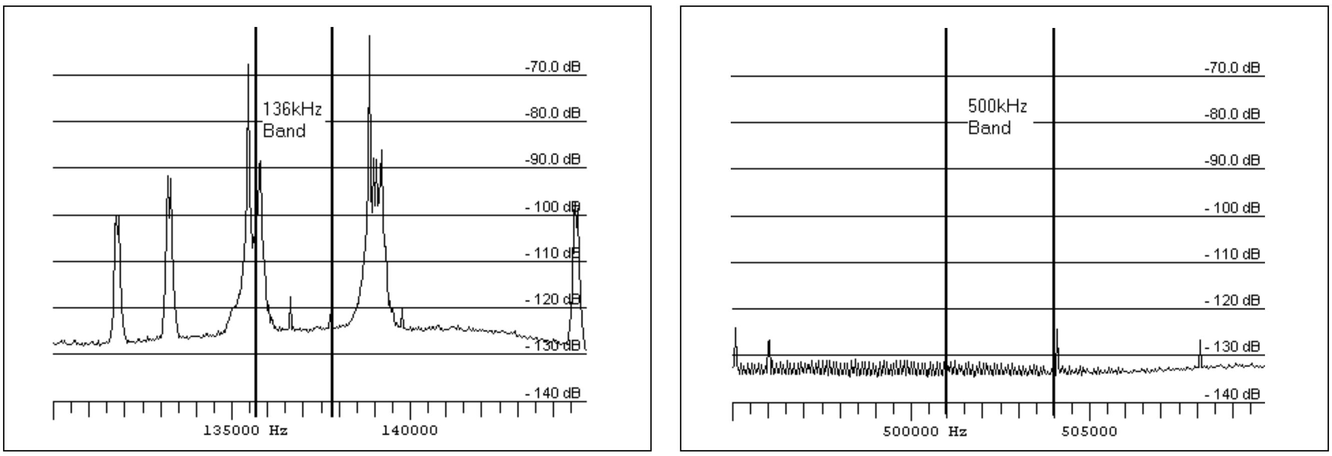


Fig 10.1: Spectrum in the vicinity of (left) 136kHz, and (right) 500kHz (vertical scale is field strength in dBV/m)

Frequency stability requirements depend on the operating mode. For CW operation, maintaining frequency within 100Hz during a contact is not usually a problem. Narrow band modes such as QRSS [2, 3] require better stability, and frequency setting accuracy. For the popular QRSS3 operating speed (widely used for 136kHz DX contacts around Europe), an initial setting accuracy, and drift of perhaps 10 - 20Hz per hour is acceptable if somewhat irritating. This level of stability can often be achieved by older receivers with mechanically tuned VFOs, provided the receiver is allowed to reach thermal equilibrium, and some means of calibrating the receiver frequency is available. These difficulties are eliminated in fully synthesised receivers, which generally exhibit a setting accuracy within 1 or 2Hz and drift a fraction of a Hertz over an extended period. This degree of frequency stability is adequate for the vast majority of applications, including the reception of inter-continental 136kHz beacon signals using QRSS30, QRSS60 and QRSS120 speeds, with bandwidths of as little as 0.01Hz, along with narrow-band, weak-signal digital modes.

For some specialised narrow band operating modes, a higher order of stability is required. This has been achieved by using high stability synthesiser reference frequency sources, such as TCXOs, OCXOs, and even atomic clock or GPS-derived frequency standards.

Amateur Receivers and Transceivers at LF

Many amateur HF receivers and transceivers can tune to 136kHz and below, and since they are already available in many shacks, probably a majority of 136kHz and 500kHz amateur stations use receivers of this type. All modern equipment is fully synthesised, so frequency accuracy and stability are good. Older receivers using multiple crystal oscillators or an interpolating VFO have relatively poor stability.

Modern crystal or mechanical CW filters have excellent shape factors, giving good rejection of strong adjacent signals. In some receivers, multiple filters can be cascaded, giving further enhanced selectivity.

Since reception at frequencies below 1.8MHz is generally included as an afterthought, manufacturers rarely specify sensitivity of amateur receivers at 136kHz and 500kHz. Unfortunately it can often be poor. There is little relation between the HF performance, cost, or sophistication of a particular model, and the sensitivity at lower frequencies. Therefore it may well be that older, cheaper models perform better at LF than their newer successors.

Few laboratory-quality sensitivity measurements are available for the LF and MF sensitivity of amateur transceivers and receivers, but the following lists some models which have been used successfully as LF receivers.

Classified as "good" are the Kenwood TS-850 (probably the most popular HF rig with 136kHz operators, but with reduced sensitivity at 500kHz - **Fig 10.2**), TS-440 and Yaesu FT-990 transceivers. Receivers include the AOR AR-7030, Icom R75, JRC NRD-345, NRD-525 and NRD-545, and Yaesu FRG-100.

Classified as "adequate" are the Icom IC-706, IC-718, IC-756Pro, IC-761, IC-765, and IC-781, Kenwood TS-140, TS-870 Yaesu FT-817 and FT-1000MP. Equipment classified as "adequate" requires either a large antenna and/or an external preamplifier to achieve adequate sensitivity. The IC-718 is fairly typical in this respect, requiring 1 microvolt at 136kHz to achieve 10dB SNR in a 250Hz bandwidth, a figure about 20dB poorer than it achieves in the HF bands.

The reason for poor sensitivity lies within the receiver front-end design. The inter-stage coupling components, in particular the first mixer input transformer, are optimised for operation at HF, and often have high losses at LF, reducing the signal level. Internally generated synthesiser noise may also be higher at LF. The front end filter used when the receiver is tuned to 500kHz or below is normally a simple low-pass filter with a cut-off frequency of 1 - 2MHz, often including an attenuator pad to reduce overloading due to medium wave broadcast signals; this further reduces sensitivity, without eliminating the broadcast signals. Some LF operators have improved receiver performance substantially by replacing the mixer input transformer with one having extended low frequency response [4]; this component must be carefully designed if receiver HF performance is to be maintained. A simpler and more common approach is to use an external preamplifier, and provide additional signal frequency selectivity, as described later in this section.

Commercial Equipment for LF

Many professional communications receivers made by such firms as Racal, Plessey, Harris, Collins, Eddystone, Rohde & Schwarz and others include coverage of the LF and MF spectrum, and surplus prices are often competitive with the amateur-type equipment discussed above. Ex-professional equipment is usually fully specified at LF and MF frequencies, so sensitivity and dynamic range are usually good at 136kHz and 500kHz. Fully synthesised professional receivers often have precision reference oscillators with excellent stability; they also usually have inputs for an external frequency reference. These features are not often found on amateur-type equipment, making them attractive if the more specialised LF communications modes are to be explored. A drawback is that affordable examples are usually fairly old, so servicing and repairs may be required from time to time. Also, they have a rather Spartan feel, with few of the 'bells and whistles' operator facilities found on modern amateur rigs. The Racal RA1792 (**Fig 10.2**) has been popular with UK amateurs on 136kHz and 500kHz. The older RA1772 also performs well.

A number of vintage receivers, including the HRO, Marconi CR100, AR88LF, cover 136kHz and 500kHz. Also, valve-era equipment designed for marine service often includes LF and MF coverage. A few amateurs have used vintage receivers for 136kHz operation. The antenna input circuit of this type of equipment is generally designed to be operated in the lower frequency ranges using an un-tuned wire antenna, and usually gives good sensitivity at 136kHz without requiring additional antenna tuners or preamplifiers. The major disadvantage of most vintage receivers is that their single-pole crystal filters have poor skirt selectivity compared to modern IF filters. This results in strong utility signals several kilohertz from the receive frequency reaching the IF and detector stages of the receiver,

Fig 10.2: Racal RA1792 (top), and RA1772 perform well on LF

causing blocking and heterodyne whistles which swamp the weak amateur signals. Unmodified vintage receivers are therefore usually poor performers at 136kHz, although for the experimentally minded, the addition of a modern IF filter and product detector could result in an effective LF CW receiver. As noted above, selectivity is less critical for 500kHz operation, and vintage receivers can perform quite well. An un-modified wartime HRO receiver has been used at M0BMU for 500kHz CW operation, with quite satisfactory results.

Selective level meters (SLMs), also called selective measuring sets or selective voltmeters, are instruments designed for measuring signal levels in the now-obsolete frequency division multiplex telephone systems; consequently, they are sometimes available surplus at low cost. SLMs are designed for precision measurement of signals down to sub-microvolt levels; their frequency range extends from a few kilohertz into the MF or HF range, so can make effective LF and MF receivers. Well-known manufacturers are Hewlett-Packard (HP3625) and the German companies Wandel and Golterman (the SPM- *selektiver pegelmesser* series; **Fig 10.3**) and Siemens.

SLMs are not purposely designed as receivers and do not have many normal receiver features, such as AGC and selectable operating modes, or sometimes even an audio output. Filter bandwidths are designed for telephony systems and are not always suited for amateur radio operating modes. Normally the 'CW pitch' is fixed at around 2kHz, so they are not well suited to CW operating, although this presents no obstacle for 'sound card' operating modes. The area where SLMs excel is in signal measurement; they have been used by a number of amateurs for 136kHz and 500kHz field strength measurements (see LF Measurements and Instrumentation section). They are often available with a tracking level generator, which is very useful for measurements on filters, or bridge-type impedance measurements.

Software-defined Radio Receivers

Software defined radio (SDR) is now becoming part of the mainstream of amateur radio, with both home constructed and commercially produced SDR hardware and software now widely available, see the chapter on software-defined radio. PC-based spectrogram software has been used for several years in conjunction with conventional receivers for the 'visual' LF/MF operating modes such as QRSS and narrow-bandwidth data modes; SDR is the natural extension of this trend.

Homebrew amateur SDR projects are most commonly based on PC-based digital signal processing software, using the PC sound card for A/D conversion of the incoming signal. Since the sound card is usually limited to 48kHz sample rate, the maximum signal frequency that can be handled by the sound card input is 24kHz. This allows direct reception of signals in the VLF range (see VLF section below), but for amateur band use, some form of external down conversion is required. This generally takes the form of an I/Q down converter, with in-phase and quadrature outputs feeding the left and right stereo inputs of the sound card.

The I/Q signal format permits image rejection to be performed by the SDR software, and also extends the bandwidth that can be processed by the sound card to 48kHz maximum; this is ample to cover the narrow amateur 136kHz and 500kHz bands with fixed, crystal-controlled conversion frequencies. All required tuning, filtering and demodulation functions are then performed in the digital domain by the SDR software. Several suites of SDR software have been made available free of charge for amateur use [5, 6, 7] that are suitable for use with I/Q down converters. This results in a very simple yet capable amateur band receiver; modifications to the well-known KB9YIG 'SoftRock' SDR receiver kits to permit 136kHz and 500kHz reception are described later in this section.

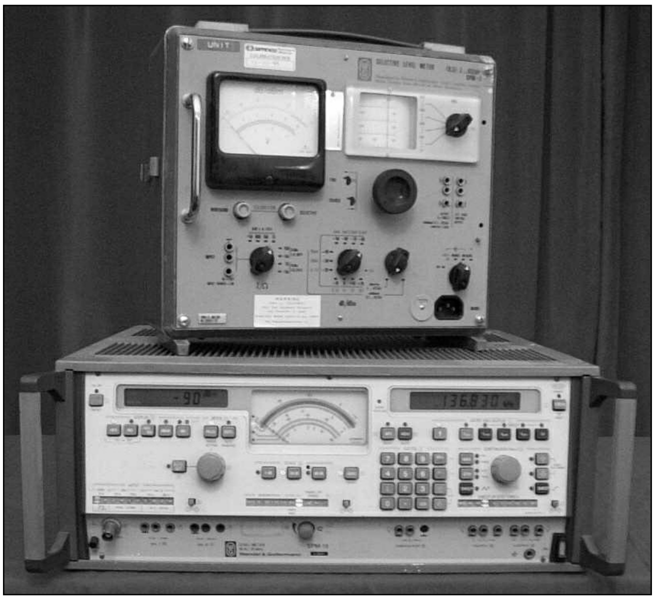


Fig 10.3: SPM-19 (bottom), and portable SPM-3 (top) selective level meters can be used as LF receivers

General coverage, direct-digitising SDR receivers are also becoming commercially available to amateurs at reasonable prices. Several amateur stations are successfully using the RFSpace Inc. SDR-IQ [8] and the Perseus SDR receiver [9] for LF and MF reception. Both these receivers are supplied with their own native SDR software, but can also be used directly with popular spectrogram software such as DL4YHF's *Spectrum Lab* [6] and *Winrad* [7] in order to generate high resolution spectrograms for 'visual modes' operation.

RECEIVE ANTENNA TUNING

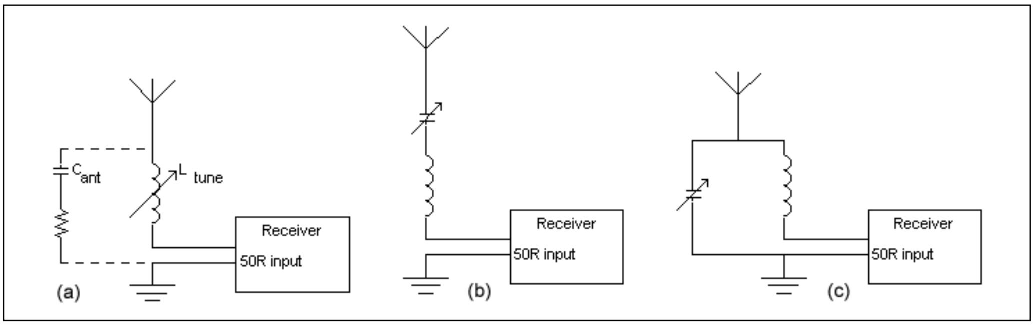
The impedance of a typical long-wire antenna at LF or MF can be modelled as a series resistor and capacitor. Taking the example in the Transmitting Antennas section of a typical long-wire antenna, the capacitance might be 287pF in series with 40 ohms at 137kHz. At 500kHz, the capacitance will be almost the same, but the resistance could be lower, perhaps 20 ohms. Assuming receiver input impedance of 50 ohms, the SWR at the feed point of the antenna is about 8200:1 at 137kHz! This mis-match results in an unacceptable signal loss of about 32dB. The loss due to mismatch at 500kHz is less severe, but still more than 20dB.

Most of the loss is caused by the capacitive reactance; signal levels can be greatly improved by resonating the antenna at the operating frequency with a series inductance. In the example above, the antenna with resonating inductance form a tuned circuit with Q around 40 and bandwidth of only a few kilohertz, which is very effective in filtering out powerful broadcast band signals.

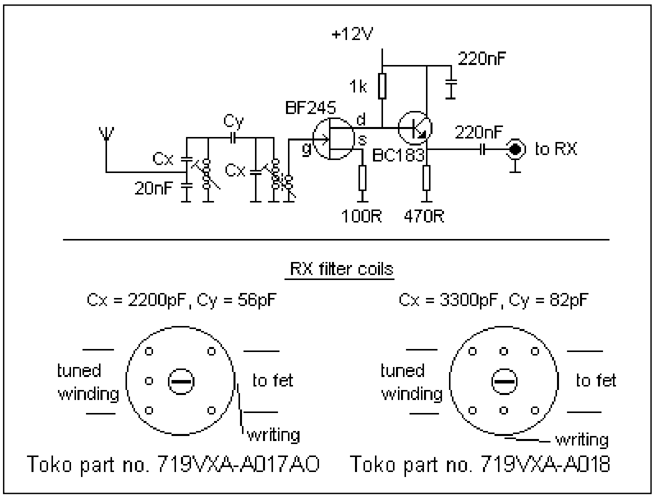
The practical effect of resonating the antenna is dramatic, normally with a long-wire antenna connected directly to the receiver, the only signals heard in the 136kHz range are numerous intermods. With the antenna tuned, these disappear and the band noise is audible above the noise floor of reasonably sensitive receivers. Attempts to receive signals at 500kHz with un-tuned wire antennas are more successful at locations where broadcast signal levels are fairly low, but an antenna tuner still yields substantial improvements.

Typical circuits used to tune wire antennas for reception are shown in **Fig 10.4**. **Fig 10.4(a)** is a simple series inductor; the value required is:

Fig 10.4: Receive antenna tuning circuits



(below) Fig 10.5: G3YXM's 136kHz preamp



$$L_{tune} = \left(\frac{1}{2\pi f \sqrt{C_{ant}}} \right)$$

with L_{tune} in henries, C_{ant} in farads and f in hertz. A useful rule of thumb is that the antenna capacitance C_{ant} will be roughly 6pF for each metre of wire, typically L_{tune} of a few millihenries will be required for 136kHz and a few hundred microhenries at 500kHz. Because of the high Q the inductance must be adjustable; this can be done using the same techniques as for transmitting antennas, or a slug-tuned coil can be used. It is often more convenient to use a fixed inductor, and adjust to resonance using a variable capacitor, as shown in **Fig 10.4(b)**. This can be a broadcast-type variable, with both sections paralleled to give about 1000pF maximum.

A higher tuning inductance is required to make up for the overall reduction in capacitance. The shunt-tuned circuit of **Fig 10.4(c)** has the convenience of one side of the tuning capacitor being grounded. The impedance match will not be quite as good, although normally perfectly adequate.

RECEIVING PREAMPS

To overcome reduced sensitivity at lower frequencies, many amateur-type receivers require a preamplifier. Because of the strong broadcast signals in the LF and MF frequency ranges, it is important that adequate selectivity is provided at the signal frequency. To obtain good S/N ratio, it is also necessary to pay attention to impedance matching between antenna and preamplifier.

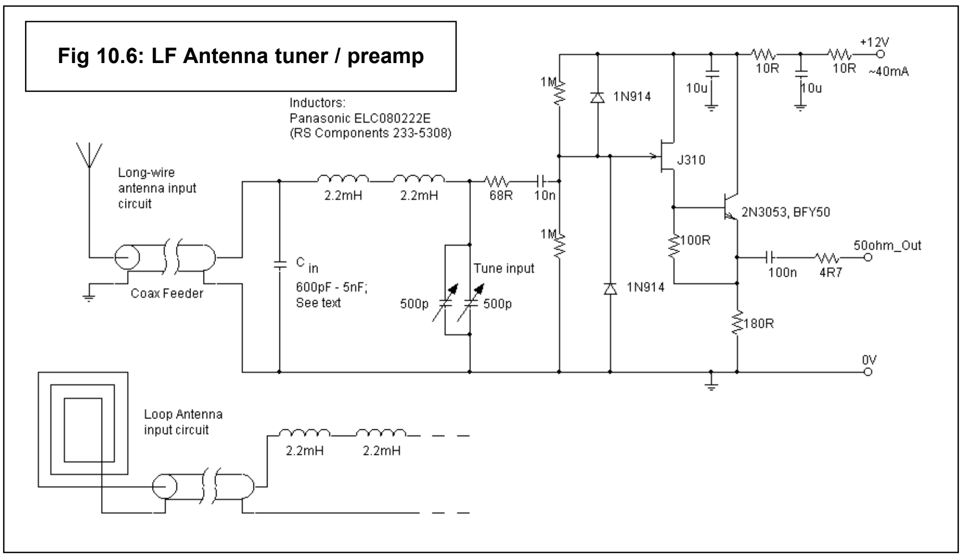
A useful and well-tried design for 136kHz due to G3YXM is shown in **Fig 10.5**, and is available as a kit with PCB from G0MRF [10]. It incorporates a double-tuned input filter which provides a bandwidth of around 3kHz centred on the amateur band. The preamp is designed for 50-ohm input impedance, so antenna matching as described in the previous section will normally be required.

The LF antenna tuner/preamp circuit of **Fig. 10.6** combines the antenna matching, filtering and preamplifier functions. It is quite flexible and can be used with a wide range of long-wire and loop antenna elements. It can easily be modified for other frequencies, including 500kHz. It has been used successfully with an IC-718 transceiver, which has fairly poor sensitivity at 136kHz. The preamp is a compound follower, with a high-impedance JFET input, and a bipolar output to drive a low impedance load with low distortion. The gain of the follower is about unity, but the high Q, peaked low-pass filter input circuit provides voltage gain, and also gives substantial attenuation of unwanted broadcast signals at higher frequencies.

The gain of the circuit depends on the type of antenna element used, of the order of 10dB with a long wire element and 30 - 40dB with a loop element. The 2.2mH inductors are the type wound on small ferrite bobbins with radial leads, and have a Q around 80 at 136kHz; other types of inductor with similar or greater Q could also be used. For wire antennas, C_{in} should be in the range 600pF - 5000pF, with large values giving a reduced signal level with longer wires, and smaller values suiting short wire antennas.

The antenna can be fed with coaxial cable, in which case the distributed capacitance of the coax (about 100pF/m for 50-ohm cable) makes up part or all of C_{in}. This allows the receiving antenna to be located remote from the shack, which is often useful in reducing interference pick-up. This circuit has given good results with wire antennas ranging from a 5m vertical

Fig 10.6: LF Antenna tuner / preamp



16 Practical VHF/UHF Antennas

Peter Swallow, G8EZE

VHF and UHF antennas differ from their HF counterparts in that the diameter of their elements are relatively thick in relationship to their length and the operating wavelength, and transmission line feeding and matching arrangements are used in place of lumped elements and ATUs.

THE (VHF) DIPOLE ANTENNA

At VHF and UHF, most antenna systems are derived from the dipole or its complement, the slot antenna. Many antennas are based on half-wave dipoles fabricated from wire or tubing. The feed point is usually placed at the centre of the dipole, for although this is not absolutely necessary, it can help prevent asymmetry in the presence of other conducting structures.

The input impedance is a function of both the dipole length and diameter. A radiator measuring exactly one half wavelength from end to end will be resonant (ie will present a purely resistive impedance) at a frequency somewhat lower than would be expected from its dimensions. Curves of 'end correction' such as Fig 16.1 show by how much a dipole should be shortened from the expected half wavelength to be resonant at the desired frequency.

The change of reactance close to half-wavelength resonance as a function of the dipole diameter is shown in Fig 16.2.

In their simplest form, dipole antennas for 2m and 70cm can be constructed from 2mm diameter enamelled copper wire and fed directly by a coaxial cable as shown in Fig 16.3. The total element length (tip to tip) should be 992mm for 145MHz operation and 326mm to cover the band 432 to 438MHz. The impedance will be around 70 ohms for most installations, so that a 50-ohm coaxial cable would present a VSWR of around 1.4:1 at the transceiver end.

A more robust construction can be achieved using tubing for the elements and moulded dipole centre boxes, available from a number of amateur radio antenna manufacturers and at radio rallies. The dipole length should be shortened in accordance with Fig 16.1 to compensate for the larger element diameters. Construction ideas and UK sources of materials can be found at [1].

Note that this simple feed may result in currents on the outside of the cable, and consequently a potential to cause interference to other electronic equipment when the antenna is used for transmitting. This can be reduced or eliminated by using a balun at the feed point.

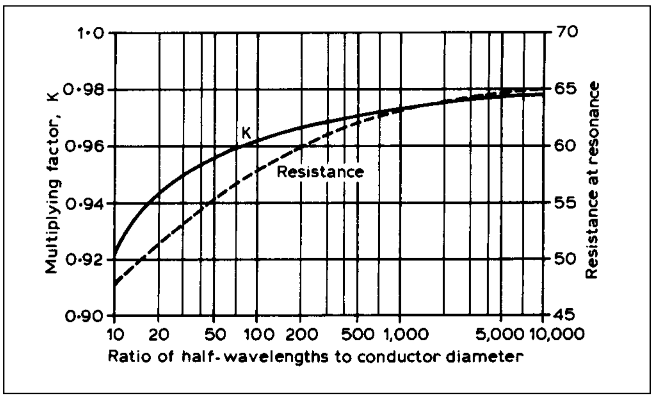


Fig 16.1: Length correction factor for half-wave dipole as a function of diameter

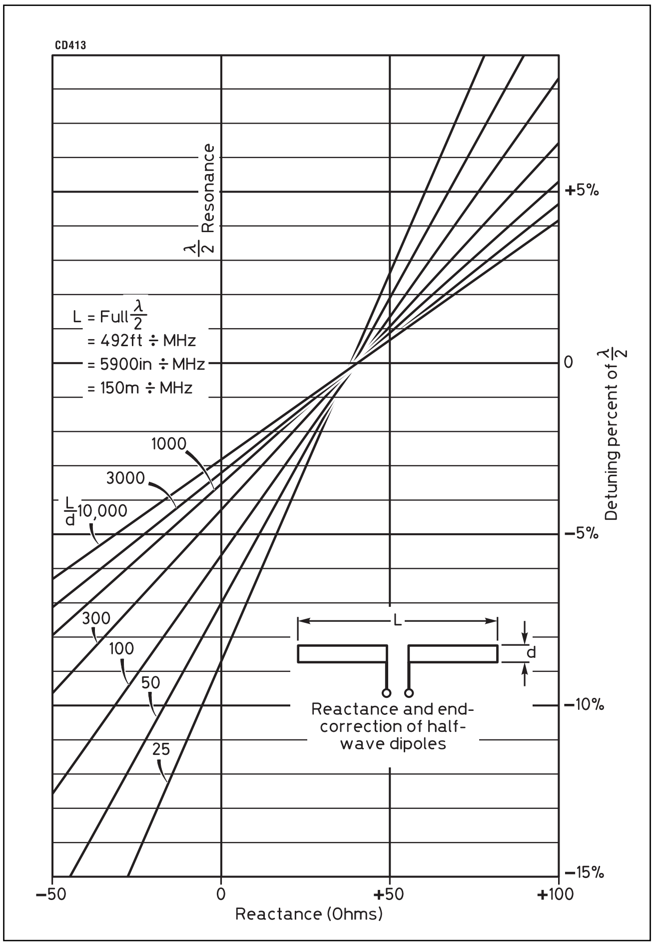


Fig 16.2: Tuning and reactance chart for half-wave dipoles as a function of diameter

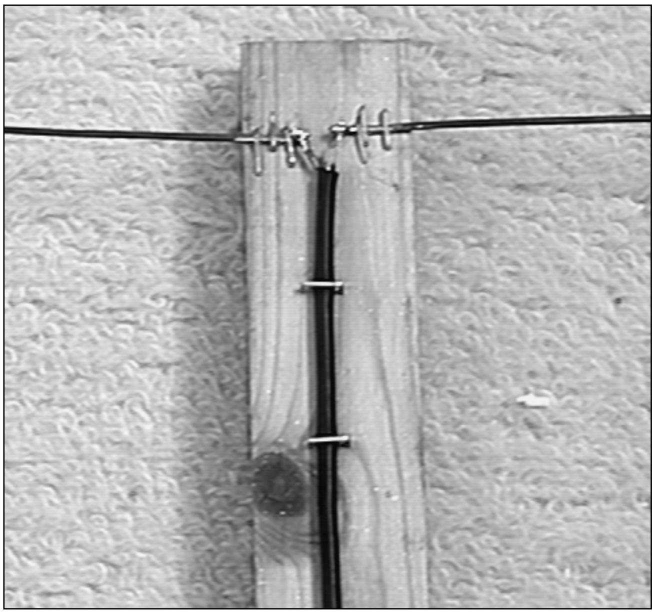


Fig 16.3: Simple dipole construction for 2m and 70cm

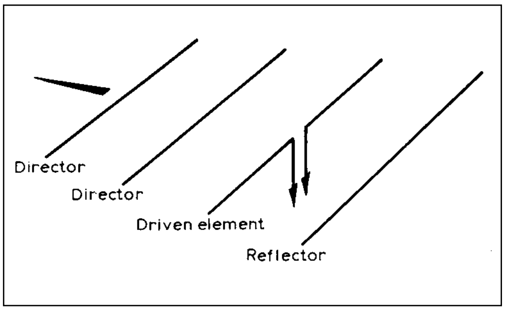


Fig 16.4: Simple Yagi antenna structure, using two directors and one reflector in conjunction with a driven element

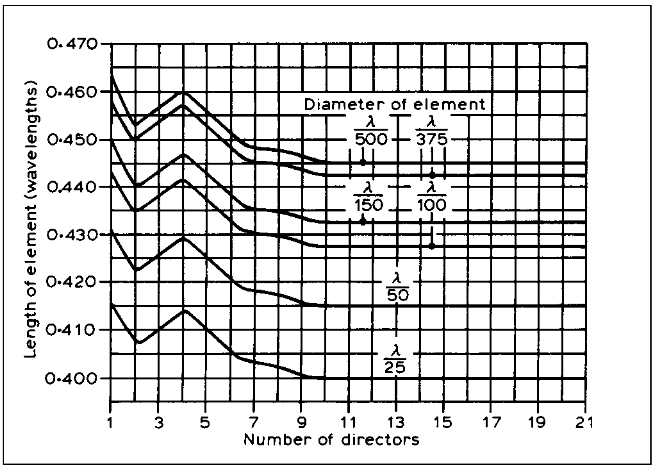


Fig 16.5: Length of director versus position in the array for various element diameters (*ARRL Antenna Book*)

THE YAGI AND ITS DERIVATIVES

The Yagi Antenna

The Yagi antenna was originally investigated by Uda and subsequently brought to Western attention by Yagi in 1928 in a form similar to that shown in **Fig 16.4**. It consists of a driven element combined with an in-line parasitic array. There have since been many variations of the basic concept, including its combination with log-periodic and backward-wave techniques.

To cover all variations of the Yagi antenna is beyond the scope of this handbook. A great number of books and many articles have been published on the subject, and a wide range of theoretical and practical pages can be found on the Internet with a simple search.

Many independent investigations of multi-element Yagi antennas have shown that the gain of a Yagi is directly proportional to the array length. There is a certain amount of latitude in the position of the elements along the array. However, the optimum resonance of each element will vary with the spacing chosen. With Greenblum's dimensions [2], in **Table 16.1**, the gain will not vary more than 1dB from the nominal value. The most critical elements are the reflector and first director as they decide the spacing for all other directors and most noticeably affect the matching. Solutions may be refined for the materials and construction methods available using one of the many software tools now freely available from the Internet, and discussed elsewhere in this handbook. These tools can be used to assess the sensitivity of a given design to alternative diameter elements and dimensions.

The optimum director lengths are normally greater the closer the particular director is to the driven element. (The increase of capacitance between elements is balanced by an increase of inductance, ie length through mutual coupling.) However, the length does not decrease uniformly with increasing distance from the driven element. **Fig 16.5** shows experimentally derived element lengths for various material diameters. Elements are mounted through a cylindrical metal boom that is two or three diameters larger than the elements.

Some variation in element lengths will occur using different materials or sizes for the support booms. This will be increasingly critical as frequency increases. The water absorbency of insulating materials will also affect the element lengths, particularly when in use, although plastics other than nylon are usually satisfactory. **Fig 16.6** shows the expected gain for various numbers of elements if the array length complies with **Fig 16.7**.

The results obtained by G8CKN using the 'centre spacing' of Greenblum's optimum dimensions shown in Table 16.1 produced identical gains to those shown in Fig 16.6. Almost identical radiation patterns (**Fig 16.8**) were obtained for both the E

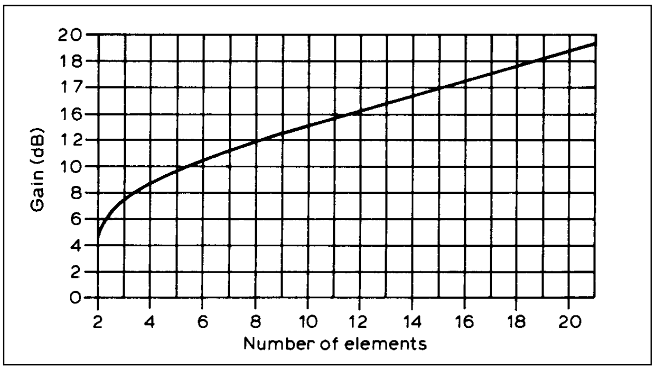


Fig 16.6: Gain in dBi versus the number of elements of the Yagi array (*ARRL Antenna Book*)

Number of elements	R-DE	DE-D1	D1-D2	D2-D3	D3-D4	D4-D5	D5-D6
2	0.15 -0.20						
2		0.07 -0.11					
3	0.16 -0.23	0.16 -0.19					
4	0.18 -0.22	0.13 -0.17	0.14 -0.18				
5	0.18 -0.22	0.14 -0.17	0.14 -0.20	0.17 -0.23			
6	0.16 -0.20	0.14 -0.17	0.16 -0.25	0.22 -0.30	0.25 -0.32		
8	0.16 -0.20	0.14 -0.16	0.18 -0.25	0.25 -0.35	0.27 -0.32	0.27 -0.33	0.30 -0.40
8 to N	0.16 -0.20	0.14 -0.16	0.18 -0.25	0.25 -0.35	0.27 -0.32	0.27 -0.32	0.35 -0.42

DE = driven element, R = reflector and D = director. N = any number. Director spacing beyond D6 should be 0.35-0.42

Table 16.1: Greenblum's optimisation for multielement Yagis

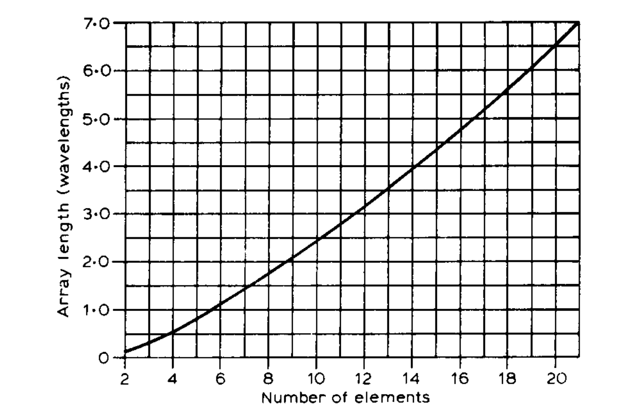


Fig 16.7: Optimum length of a Yagi antenna as a function of the number of elements (*ARRL Antenna Book*)

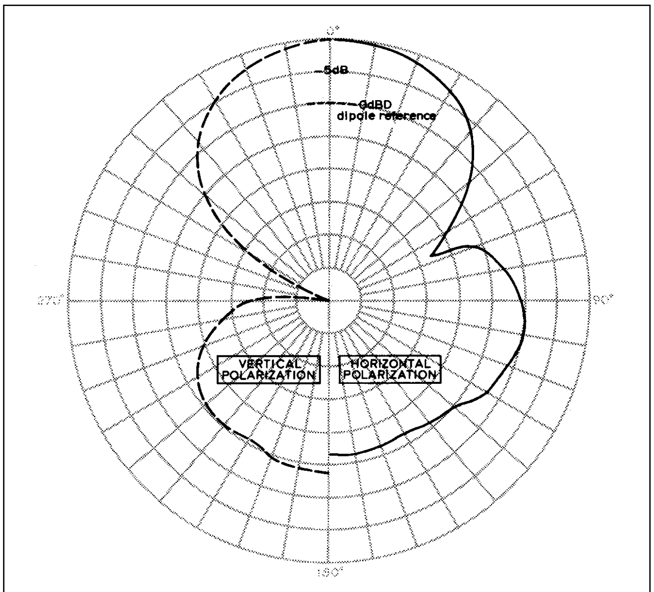


Fig 16.8: Radiation pattern for a four-element Yagi using Greenblum's dimensions

and H planes (V or H polarisation). Sidelobes were at a minimum and a fair front-to-back ratio was obtained.

Considerable work has been carried out by Chen and Cheng on the optimising of Yagis by varying both the spacing and resonant lengths of the elements [3].

Table 16.2 and **Table 16.3** show some of their results obtained in 1974, by optimising both spacing and resonant lengths of elements in a six element array.

Table 16.3 shows comparative gain of a six element array with conventional shortening of the elements or varying the element lengths alone. The gain figure produced using conventional shortening formulas was 8.77dB relative to a λ/2 dipole (dBd). Optimising the element lengths produced a forward gain of 10dBd. Returning to the original element lengths and optimising the element spacing produced a forward gain of 10.68dBd. This is identical to the gain shown for a six-element Yagi in Fig 16.6. Using a combination of spacing and element length adjustment obtained a further 0.57dBd, giving 11.25dBd as the final forward gain as shown in Table 16.3.

A publication of the US Department of Commerce and National Bureau of Standards [4], provides very detailed experimental information on Yagi dimensions. Results were obtained from measurements to optimise designs at 400MHz using a model antenna range.

The information, presented largely in graphical form, shows very clearly the effect of different antenna parameters on realisable gain. For example, it shows the extra gain that can be achieved by optimising the lengths of the different directors, rather than making them all of uniform length. It also shows just what extra gain can be achieved by stacking two elements, or from a 'two-over-two' array.

The paper presents:

(a) The effect of reflector spacing on the gain of a dipole.
(b) Effect of different equal-length directors, their spacing and number on realisable gain.
(c) Effect of different diameters and lengths of directors on realisable gain.
(d) Effect of the size of a supporting boom on the optimum length of parasitic elements.
(e) Effect of spacing and stacking of antennas on gain.
(f) The difference in measured radiation patterns for various Yagi configurations.

	h_1/λ	h_2/λ	h_3/λ	h_4/λ	h_5/λ	h_6/λ	Gain (dBd)
Initial array	0.255	0.245	0.215	0.215	0.215	0.215	8.78
Length-perturbed array	0.236	0.228	0.219	0.222	0.216	0.202	10.00

$b_i1 = 0.250\lambda$, $b_i2 = 0.310\lambda$ (i = 3, 4, 5, 6), a = 0.003369λ

Table 16.2: Optimisation of six-element Yagi-Uda array (perturbation of element lengths)

	h1/λ	h2/λ	h3/λ	h4/λ	h5/λ	h6/λ	b21/λ	b22/λ	b43/λ	b34/λ	b35/λ	Gain (dBd)
Initial array	0.255	0.245	0.215	0.215	0.215	0.215	0.250	0.310	0.310	0.310	0.310	8.78
Array after spacing perturbation	0.255	0.245	0.215	0.215	0.215	0.215	0.250	0.289	0.406	0.323	0.422	10.68
Optimum array after spacing and length perturbations	0.238	0.226	0.218	0.215	0.217	0.215	0.250	0.289	0.406	0.323	0.422	11.26

Table 16.3: Optimisation for six-element Yagi-Uda array (perturbation of element spacings and element lengths)

Length of Yagi (λ)	0.4	0.8	1.20	2.2	3.2	4.2
Length of reflector (λ)	0.482	0.482	0.482	0.482	0.482	0.475
Length of directors (λ):						
1st	0.424	0.428	0.428	0.432	0.428	0.424
2nd	-	0.424	0.420	0.415	0.420	0.424
3rd	-	0.428	0.420	0.407	0.407	0.420
4th	-	-	0.428	0.398	0.398	0.407
5th	-	-	-	0.390	0.394	0.403
6th	-	-	-	0.390	0.390	0.398
7th	-	-	-	0.390	0.386	0.394
8th	-	-	-	0.390	0.386	0.390
9th	-	-	-	0.398	0.386	0.390
10th	-	-	-	0.407	0.386	0.390
11th	-	-	-	-	0.386	0.390
12th	-	-	-	-	0.386	0.390
13th	-	-	-	-	0.386	0.390
14th	-	-	-	-	0.386	-
15th	-	-	-	-	0.386	-
Director spacing (λ)	0.20	0.20	0.25	0.20	0.20	0.308
Gain (dBd)	7.1	9.2	10.2	12.25	13.4	14.2

Element diameter 0.0085λ. Reflector spaced 0.2λ behind driven element. Measurements are for 400MHz by P P Viezbicke.

Table 16.4: Optimised lengths of parasitic elements for Yagi antennas of six different boom lengths

The highest gain reported for a single boom structure is 14.2dBd for a 15-element array (4.2λ long and reflector spaced at 0.2λ, with 13 graduated directors). See **Table 16.4**.

It has been found that array length is of greater importance than the number of elements, within the limit of a maximum element spacing of just over 0.4λ.

Reflector spacing and, to a lesser degree, the first director position affects the matching of the Yagi. Optimum tuning of the elements, and therefore gain and pattern shape, varies with different element spacing.

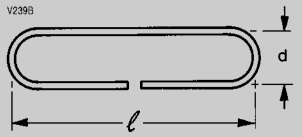
Near-optimum patterns and gain can be obtained using Greenblum's dimensions for up to six elements. Good results for a Yagi in excess of six elements can still be obtained where ground reflections need to be minimised.

Chen and Cheng employed what is commonly called the long Yagi technique. Yagis with more than six elements start to show an improvement in gain with fewer elements for a given boom length when this technique is used.

As greater computing power has become available, it has been possible to investigate the optimisation of Yagi antenna gain more extensively, taking into account the effects of mounting the elements on both dielectric and metallic booms, and the effects of tapering the elements at lower frequencies.

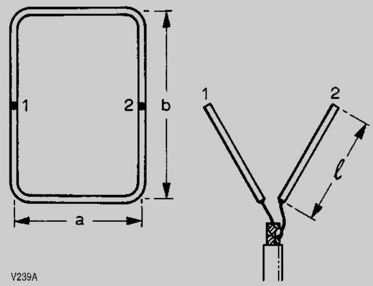
Dr J Lawson, W2PV, carried out an extensive series of calculations and parametric analyses, collated in reference [5], which although specifically addressing HF Yagi design, explain many of the disappointing results achieved by constructors at VHF and above. In particular, the extreme sensitivity of some designs to minor variations of element length or position are revealed in a series of graphs which enable the interested constructor to select designs that will be readily realisable.

	70.3MHz	Length 145MHz	433MHz
Driven elements			
Dipole (for use with gamma match)	79 (2000)	38 (960)	12 3/4 (320)
Diameter range for length given	1/2 - 3/4 1/8 - 1/4 (12.7 - 19.0)	(6.35 - 9.5)	1/4 - 3/8 (3.17 - 6.35)



	70.3MHz	145MHz	433MHz
Folded dipole 70-ohm feed			
l length centre-centre	77 1/2 (1970)	38 1/2 (980)	12 1/2 (318)
d spacing centre-centre	2 1/2 (64)	7/8 (22)	1/2 (13)
Diameter of element	1/2 (12.7)	1/4 (6.35)	1/8 (3.17)

a



	70.3MHz	145MHz	433MHz
centre/centre	32 (810)	15 (390)	5 1/8 (132)
b centre/centre	96 (2440)	46 (1180)	152 (395)
Delta feed sections (length for 70Ω feed)	22½ (570)	12 (300)	42 (110)
Diameter of slot and delta feed material	1/4 (6.35)	3/8 (9.5)	3/8 (9.5)
Parasitic elements			
Element			
Reflector	85 1/2 (2170)	40 (1010)	13 1/4 (337)
Director D1	74 (1880)	35 1/2 (902)	11 1/4 (286)
Director D2	73 (1854)	35 1/4 (895)	11 1/8 (282)
Director D3	72 (1830)	35 (890)	11 (279)
Succeeding directors	1in less (25)	1/2in less (13)	1/8in less
Final director	2in less (50)	1in less (25)	3/4in less
One wavelength (for reference)	168 3/4 (4286)	81 1/2 (2069)	27 1/4 (693)
Diameter range for length given	1/2 - 3/4 (12.7 - 19.0)	1/4 - 3/8 (6.35 - 9.5)	1/8-¾ (3.17 - 6.35)
Spacing between elements			
Reflector to radiator	22 1/2 (572)	17 1/2 (445)	5 1/2 (140)
Radiator to D1	29 (737)	17 1/2 (445)	5 1/2 (140)
D1 to D2	29 (737)	17 1/2 (445)	7 (178)
D2 to D3, etc	29 (737)	17 1/2 (445)	7 (178)

Dimensions are in inches with millimetre equivalents in brackets.

Table 16.5: Typical dimensions of Yagi antenna components. Dimensions are in inches with metric equivalents in brackets

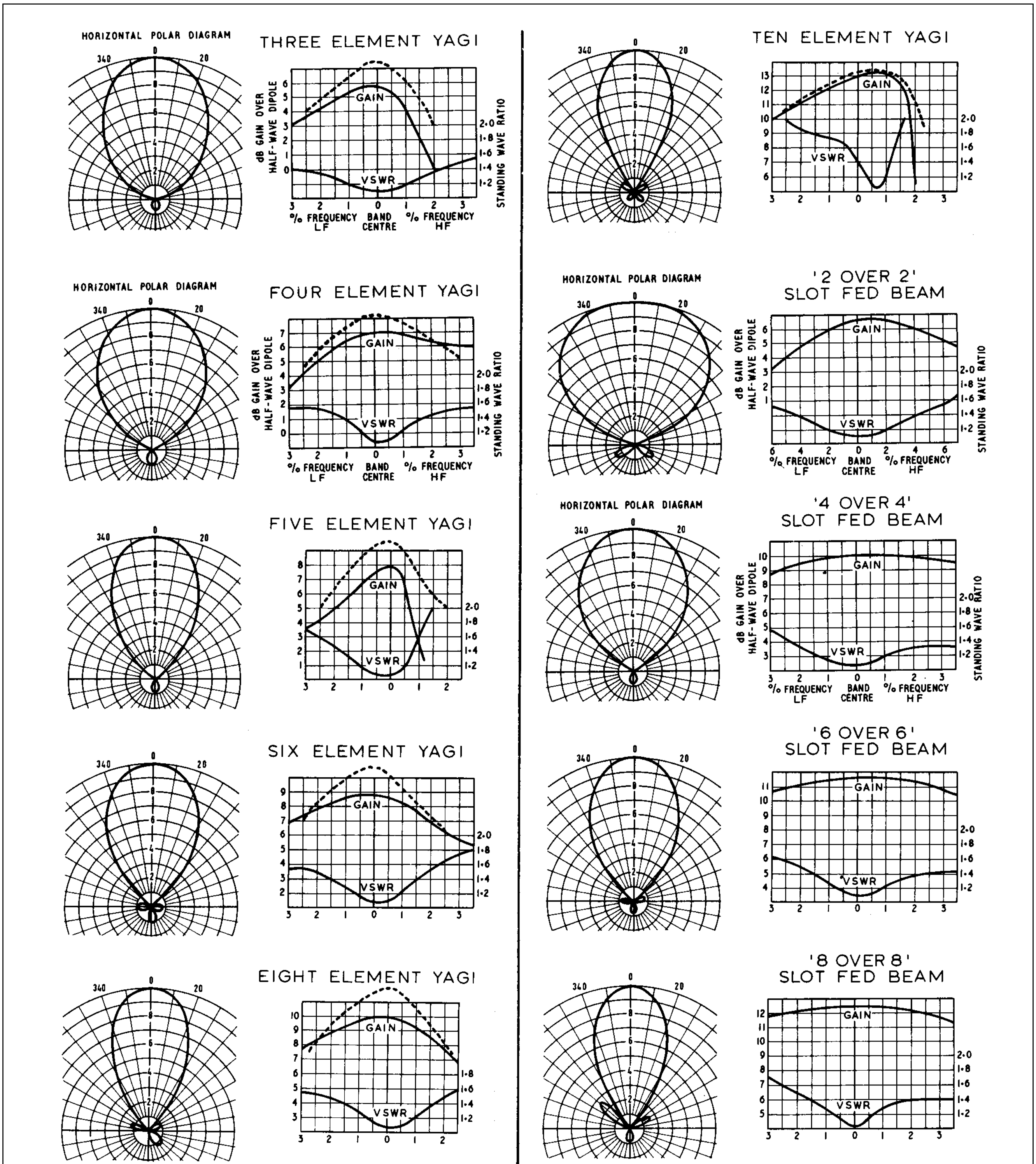


Fig 16.9: Charts showing voltage polar diagram and gain against VSWR of Yagi and skeleton-slot antennas. In the case of the six Yagi antennas the solid line is for conventional dimensions and the dotted lines for optimised results discussed in the text

The keen constructor with a personal computer may now also take advantage of modelling tools specifically designed for optimisation of Yagi antennas and arrays, eg [6], although some care is needed in their use if meaningful results are to be assured. The Internet is a good source for Yagi antenna design and optimisation programmes, many of which can be obtained free of charge, or for a nominal sum.

From the foregoing, it can be seen that several techniques can be used to optimise the gain of Yagi antennas. In some circumstances, minimisation of sidelobes is more important than maximum gain, and a different set of element spacings and lengths would be required to achieve this. Optimisation with so many independent variables is difficult, even with powerful computing methods, as there may be many solutions that yield comparable results.

Techniques of 'genetic optimisation' have been developed and widely adopted, which can result in surprising, but viable designs [7], [8]. The technique requires the use of proven computer-based analysis tools such as NEC, MININEC or their derivatives.