Quad comparative amplifier tests

Tuesday, 21st March, 1978

It is our belief that it is perfectly possible to design amplifiers that do not degrade the quality of the programme. By this we mean that if the programme is monitored and aurally compared before and after it has passed through the amplifier in question, there will be no detectable change in quality. We further believe that if an amplifier does not meet this criterion then quite rational objective reasons can be advanced to account for the discrepancy, although this may involve parameters not normally incorporated in published specifications. It is possible to devise a series of subjective tests which confirm these beliefs with a high degree of probability.

All this is not to say that all amplifiers are the same. There is a considerable sprinkling of amplifiers—in all price ranges—which do change (distort) the quality for good or ill. This has led to comparisons of amplifiers by listening tests in various parts of the world which purport to rank amplifiers by simple substitution, a seemingly quite straight forward procedure. The fact that this rank order varies widely from test to test and from place to place, leads to the suspicion that perhaps other errors are masking the true situation.

Investigations which we and others (1) (2) (3) have carried out indicate that technical errors in the 'comparator' systems used, together with errors of a psychological nature due to the test procedure are indeed sufficient to account for the anomalous results.

The purpose of the present test series is to ascertain whether it is possible to reduce these errors to a sufficient degree so that they do not unduly mask the true findings.

It is a further purpose of the test to put into perspective the differences, if any other than power output, of amplifiers made by this company over a quarter of a century.

We therefore commissioned James Moir & Associates to organise and carry out panel listening tests in which the only variables would be the three amplifiers type Quad II, 303 and 405. It may be of interest to note that these three designs cover very different technologies and between them allow comparisons of—

- Valves *v.* transistors
- Class A *v.* Class B
- Output condenser *v.* direct coupled
- Separate *v.* common power supplies
- Regulated *v.* unregulated power supplies
- Transformer *v.* transformerless output
- Sharp cut off at 20Hz *v.* flat to near DC.

1 Technical Considerations

1(a) Choice of Loudspeaker

One of the proposed panel members had very strong views that all amplifiers sound different and the choice or loudspeaker Yamaha NS1000 was largely to meet his wishes, though with the agreement of other members of the panel. His subsequent withdrawal from these tests occurred too late to reconsider the loudspeaker choice because all the necessary impedance measurements had already been made on this load. No other inference should be drawn from the choice of loudspeaker.

The choice of loudspeaker can have two effects on the outcome of the tests, (a) the variation in the ability of different loudspeakers to reveal subtle effects in different areas and (b) different impedance curves will introduce different but similar magnitude errors in response (see source impedance). The resistive and reactive components of the loudspeaker cables are negligible for the loudspeaker used.

1(b) Source Impedance

Apart from power output the principle difference between the three amplifiers is in their output source impedance. The effect of this can best be seen by plotting the frequency response when driving the actual loudspeaker load. These curves are attached. It has been argued that since this is to a first order a linear deviation it should be corrected. We have not done this as it would depart from the spirit of the test. These frequency deviations should be extremely difficult to detect aurally although it is necessary to adjust the relative level between the amplifiers so that the deviations fluctuate about a mean, taking variations of ear sensitivity with frequency into account. The levels are set at 2000Hz as shown by the curves.

1(c) Multiple Quad II's

It is highly desirable to adjust the volume level in subjective tests to suit the wishes of the listening panel. A trial run with the same loudspeakers in the same room (but with a different panel) gave a demand on peaks of 20V and 3·25A. This is beyond the capability of a single Quad II amplifier. Fortunately it is possible, by linking the output transformer of each Quad II for 16Ω and paralleling three outputs, to obtain a resultant source impedance for both resistive and reactive components which is the same as a single Quad II linked for 8Ω. Thus the performance is identical except for increased voltage and current capability before overload.

The output capability of the amplifiers in terms of peak volts and current is—

Compound Quad II		22V	4·2A
	303	30V	3·5A
	405	40V	3·5–7A (voltage dependent)
Single Quad II for comparison		15V	2A

This makes it possible to produce adequately high sound levels without overloading the amplifiers, an essential requirement for any meaningful test.

1(d) Electrical Interaction

It is highly important that each amplifier is considered as a four terminal network, i.e. that the input 'earth' and output 'earth' have no common connections other than in the amplifiers themselves. The effect of any slight amount of input/output coupling (e.g. as little as 10 milliohms shared between output and input current) will affect each amplifier depending upon its voltage gain and whether or not it is inverting. The voltage gains of the amplifiers are:

Quad II	11 to 1 non-inverting
Quad 303	42 to 1 inverting
Quad 405	56 to 1 inverting

The most sensitive method of ascertaining the presence or otherwise of voltage or current coupling is to measure the amplifier source impedance throughout the frequency range with the amplifiers connected in the test configuration. Indeed a proper check of other amplifier parameters in the test configuration is mandatory.

The equivalent source impedance components of the three systems measured in situ at the junction between the test set up and the input to the loudspeaker cables are:

Quad 405	$\cdot 23\Omega + 5\cdot 3\mu H$
Quad 303	$\cdot 47\Omega + 8\mu H + 2200\Omega F$
Compound Quad II	$\cdot 7\Omega + 10\mu H$

The output of the amplifiers are completely isolated and fed through double-pole relays having gold plated contacts to the loudspeakers and the connections are such that signal polarity is preserved regardless of the inversion of two of the three systems.

The inputs to each amplifier should not be fed from impedances greater than that for which they are designed.

The inputs are each connected to potentiometers of the correct impedance potting down from a source of low impedance direct from a Studer A80 professional tape recorder. The inputs to the Quad II's are independently potted and balanced to maintain current sharing within the required tolerance.

2 Test Conditions

2(a) Test Tapes

These were 15in/sec (38cm/sec) first generation copies of music specially selected from a large quantity of material available from several of the premier recording studios in this country, and from other sources. These samples were further distilled by careful listening comparisons until we were left with four selections that were considered to be outstanding in respect of frequency response, low distortion and acoustic clarity. The examples of programme finally used consisted of a concert orchestra, a light orchestral section, a group of male singers and finally a 'pop' group, all thought to be broadly representative of the music played at home by the average enthusiast.

2(b) Loudspeakers

The loudspeakers were mounted on 260mm high stands at a distance of 700mm from the wall behind the speaker and 700mm from the side wall. The speakers were then 2·5m apart. The units were used without the front cloth.

Catalogue data for these speakers is included but frequency response and distortion data were measured on the units actually used. The frequency response was good without being outstanding, but the amplitude distortions were particularly low. The impedance data shows that the minimum impedance is reached around 100Hz and is approximated 5 ohms.

2(c) Room Acoustics

The listening room has dimensions of $3\cdot 9 \times 5\cdot 5 \times 2\cdot 4$m and a volume of $51\cdot 5m^3$ and is furnished in the conventional manner as a lounge. The reverberation time/frequency relation of the room with six people seated is 400mS seconds at 500Hz, falling to 300mS seconds at 4Hz, but complete data is shown in the Appendix.

3 Presentation

3(a) The Listening Panel

The six members of the listening panel were selected to include those with a high reputation for the judgement of sound quality. They were —

Laurie Fincham, B.Sc.(Eng)	*Director and Chief Engineer, KEF Electronics Ltd.*
John Crabbe	*Editor, Hi-Fi News & Record Review*
Mike Ballance	*Deputy Editor, Popular Hi-Fi*
Jim Rogers	*Technical and Research Director, J. R. Loudspeakers Ltd.*
David Stripp	*Assistant to Chief Engineer B.B.C. Radio, responsible for Sound Quality*
John Borwick, B.Sc.	*Audio Editor of 'Gramophone' and Senior Lecturer in Recording Techniques, Department of Music, University of Surrey*

The panel sat in two rows at a mean distance of 4m from the loudspeakers.

3(b) Procedure

The three amplifiers were presented in random order to allow paired comparison tests between all the possible combinations of two amplifiers. Each member of the panel indicated his preference, or lack of preference, for the two amplifiers being compared, but every pair was identified only as A or B. The panel members could not know the type of amplifier to which they were listening at any particular time, although they were told in the preamble that on occasions both A and B could be the same amplifier. The test consisted of presenting a 30-second selection of programme on 'A' amplifier, followed by a repeat of the same section on 'B' amplifier, with a gap of a few seconds between the two pieces.

At the end of the repeat the panel members listed their decision on a score sheet in the form of:

I prefer 'A'
I prefer 'B'
I have no preference.

Though forming no part of the statistical analysis the panel were asked to indicate their reason for any expressed preference.

The question was phrased in terms of 'preference' rather than 'difference', as this forced choice gives more statistical information without significantly weakening the absolute determination of difference, and in less than 5% of cases was a difference noted with no preference stated.

The score sheets were collected after each series of six tests, of all four musical excerpts, twenty-four series of comparisons being made. Each group of six tests occupied an hour including an interval of 15–20 minutes in the middle. The panel members were asked to refrain from discussing the results before or after marking up their score cards. Two series of six tests were run during the morning and a further two series of tests run during the afternoon.

This test procedure closely follows the recommendations of the International Electrotechnical Commission—I.E.C.

4 Results

The data summarised from the score sheets is attached. There is sufficient data to allow two of the standard statistical tests to be applied to determine how far the result obtained is likely to be due to sheer chance (luck) rather than by any real difference in the performance of the three amplifiers. The 50% probability test applied to a paired comparison of samples reveals how far the result obtained is due to sheer chance and how far it is due to a real difference.

As a second test of the validity of the listening panel's opinion the Chi^2 test was applied to their scoring. Both tests confirm that the preferences expressed by the panel were no more than would be achieved by sheer chance.

It is worthy of comment that a week before the listening tests described we carried out a trial run to check the operation of the switching system and the scoring and analysis techniques. This used a different but equally expert listening panel and a different selection of musical programme, but analysis of the data obtained showed that the consensus decision of this trial panel was in excellent agreement with the findings of the second panel, and confirmed that the decisions of the panel were no better than might be expected from sheer chance.

References referred to on page 1

1 Audible amplifier distortions is not a mystery.
Peter J. Baxandall
Wireless World, November 1977.

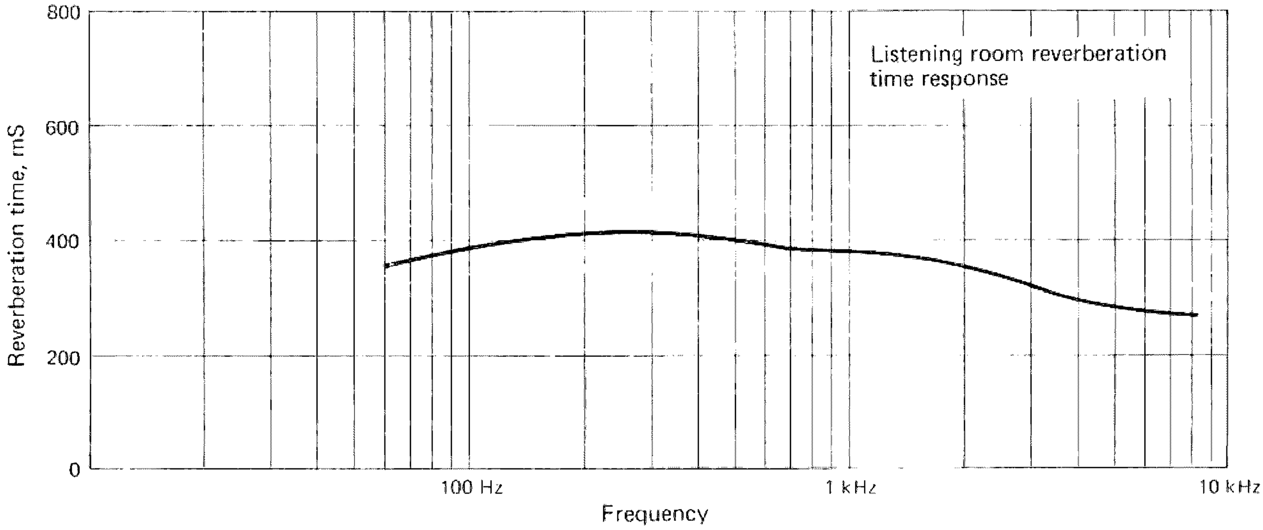
2 Six amplifiers—how did they sound?
Ian G. Masters
Audio Scene—Canada, May 1977.

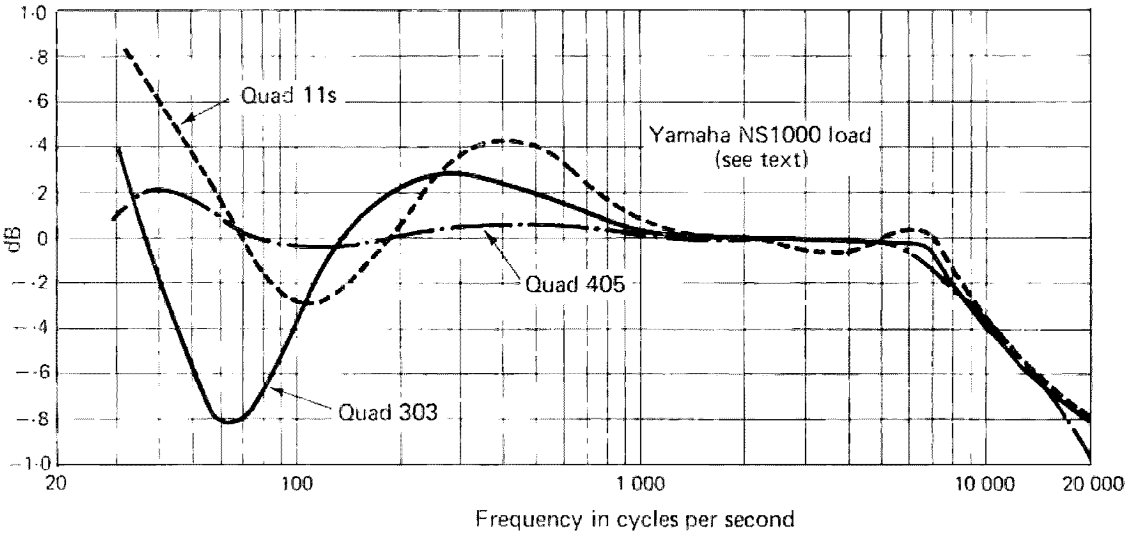
3 Rational Amplifier Testing
Peter J. Walker
Positive Feedback
Hi-Fi News & Record Review, July 1977.

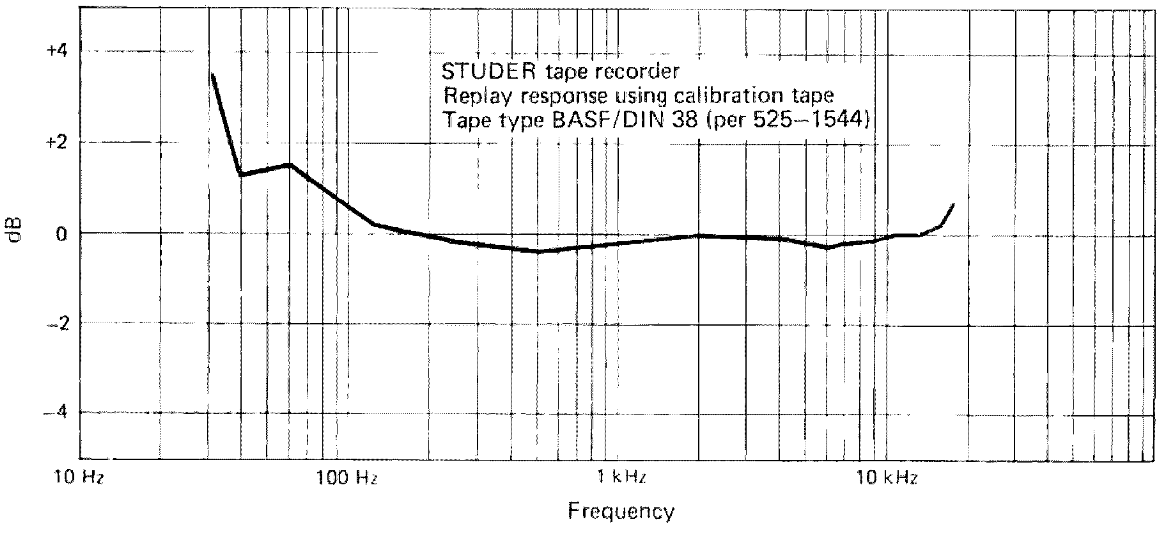
Statistical Assessment

In a paired comparison test of two identical systems a large number of comparisons must be made if the identity is to be confirmed. Thus if an unbiassed penny is thrown three times it will necessarily come down one way twice as often as it comes down the other way, but as the number of throws is increased the number of heads and tails will tend towards equality. The same principles apply to a comparison of the sound quality from two identical or very similar amplifiers. A very considerable number of comparisons are required to confirm the relative performance of the two amplifiers. Any small number of comparisons is likely to show a judgement in favour of either amplifier by sheer chance alone. If a very large number of comparisons cannot be made for good practical reasons then it is necessary to apply an appropriate mathematical technique to the analysis of the voting to determine the probability of the consensus opinion being due to chance alone.

Many techniques have been developed for assessing the probability of any given result being







due to chance. These methods are described in many statistical texts. From these available techniques the test known as the '50% Probability Test' appears appropriate for testing the result of a series of paired comparison tests.

The judging panel are asked to record their view in the following alternative forms:

1 I prefer system A.
2 I prefer system B.
3 I have no preference.

The process of estimating how far the panel decision is due only to chance is then as follows. The total number of preferences for the 'A' system is totalled, as are the number of preferences for the 'B' system. The smallest of these totals 'X' is added to the larger total 'Y' to obtain the total number 'N' of expressed preferences.

Standard statistical Tables relating 'X' and 'N' are then consulted to obtain the probability that any observed preference is the result of sheer chance. If the probability is found to be less than 10% then it is likely that the observed result is the right result and as the probability moves towards 1% the observed result is increasingly likely to be the correct result. If the probability is found to be higher than 10% the observed result is increasingly likely to be the result of mere chance. In the series of comparisons dealt with in this report the probability that the observed result was due to chance was always well above 10% indicating that where preference was indicated by the voting it was almost certainly due to chance and had little connection with any real differences.

Chi^2 Test

A chi-squared test has been applied to the results to study the difference between the percentage of 'No Preference' votes cast on identical amplifier comparisons and non identical amplifier comparisons. It could be argued that the percentage of 'No Preference' votes chosen on identical amplifier comparisons should be quite high (greater than 90%) if the panel were reasonably competent but this is felt to be quite unfair as the listening panel were seriously attempting to draw distinctions between amplifiers which they knew were near identical in terms of technical performance.

To apply the test we assume (possibly incorrectly) that there is no audible difference between the amplifiers tested and that the panel are guessing. On the identical amplifier comparisons we then calculate (as a percentage) the number of times each individual cast a 'No Preference' vote. For example, out of 24 identical amplifier comparisons, listener (d) cast 16 'No Preference' votes (i.e. 66·66%). Thus if we assume that listener (d) maintains this standard of voting we should expect the same percentage of 'No Preference' votes to be cast when different amplifiers are compared.

The results were as follows:

TABLE

	% 'No Preference' votes for identical amplifiers	*Observed 'No Preference' votes on dissimilar amplifiers	Expected 'No Preference' votes on dissimilar amplifiers
(a)	54·16%	39	39
(b)	87·50%	57	63
(c)	70·80%	50	50·97
(d)	66·66%	42	47·99
(e)	87·50%	59	63
(f)	41·60%	28	29·95

If we now apply a chi-squared test, i.e.

$$\text{Chi}^2\ (\chi^2) = \sum \frac{(O - E)^2}{E} \text{ which for the above figures} = 1{\cdot}72$$

where O = observed value

E = expected value

NB* 72 votes are possible.

If the observed and expected results are near identical the panel cannot have been listening to any real difference between the amplifiers, the results obtained probably being the same even if only one amplifier had been used throughout the tests. If however, the observed and expected results are significantly different it is conclusive proof that some audible difference must exist between the amplifiers.

The null hypothesis used on the test results is that no difference exists between the observed 'No Preference' and the expected 'No Preference' votes. With reference to the Chi² tables with a system having five degrees of freedom the 5% Pr level = 11·07. Our value is significantly less than this, so we accept the null hypothesis that there is no significant difference between the results.

A third test can in fact be applied which measures the degree of correlation between the number of 'No Preference' votes cast on identical amplifier comparisons and the number cast on dissimilar amplifier comparisons. The sample correlation coefficient r is given by:

$$r = \frac{\Sigma xy - \dfrac{\Sigma x^2 y}{n}}{\sqrt{\left(\Sigma x^2 - \dfrac{(\Sigma x)^2}{n}\right)\left(\Sigma y^2 - \dfrac{(\Sigma y)^2}{n}\right)}}$$

With a result giving equal or greater than 0·9 the correlation is said to be good, whilst r lower than 0·5 the correlation is said to be poor or non-existent. (A high negative value of r would indicate good negative correlation.) Without going into the mathematics it is sufficient to say that the results found substantially confirm the conclusion of the chi² test, the correlation coefficient being greater than 0·98.

A second complete set of data was available for analysis from an earlier comparison trial, the tests being carried out to ensure that the switching and amplifier levels etc., were in order for the official tests. It is relevant to note that the results from there earlier tests substantially confirm the conclusions reached using the official test data.

Technical specification

Loudspeakers

Yamaha NS1000M monitor class loudspeakers were used throughout the tests, the manufacturers specification being as follows:

Height	675mm
Width	375mm
Depth	326mm
Weight	31kG
Crossover frequencies	500Hz and 6kHz
Frequency response	40Hz–20kHz
Impedance	8 ohms nominal (5ohms at 100Hz)
Recommended Maximum Amplifier power	100 watts RMS
Sensitivity for 2·82V pink noise at 1 metre	84·5dBA

Quad 405 specifications

Measurements apply to either channel, with or without the other channel operating.

Power Output

The amplifier is intended for use with loudspeakers of 4–16Ω nominal impedance.

Power and distortion for various frequencies. Continuous sine wave into 8Ω resistive load.

100Hz any level up to 100 watts
 <0·01% Dtot

1kHz any level up to 100 watts
 <0·01% Dtot

10kHz any level up to 100 watts
 <0·05% Dtot

For other impedances and frequencies see graphs.

Notes: In addition to the performance into a resistive load R, the amplifier will maintain full voltage within the same distortion rating into a load R + jX where X is any value from zero to infinity.

With the additional power limiter inserted the maximum output voltage is limited to 20V rms ±10% (50 watts 8Ω) all other performance figures unchanged.

Output Internal Impedance and Offset

3·3μH in series with 0·03Ω. Offset <7mV.

Frequency Response

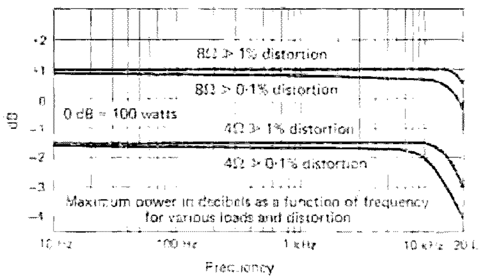
Ref. 1kHz.

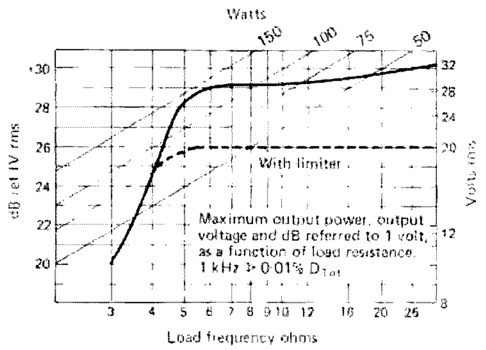
Low frequency dB at 20Hz

Filter attenuation as curve.

High frequency 0·5dB at 20kHz

3dB at 50kHz



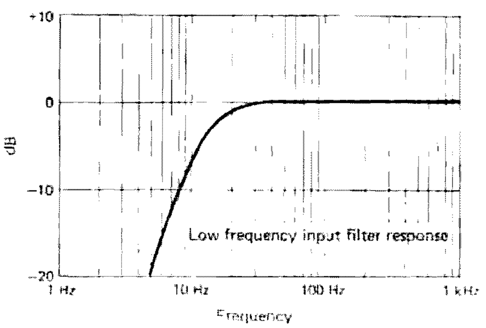


Signal Input Level
0·5V rms ± 0·5dB for 100 watts into 8Ω. Amplifier loads and input by 20kΩ in parallel with 50pF.

Signal Input Slew Rate Limit
0·1V/μS.

Provided the rate of change of input voltage does not exceed this figure and the amplifier is not driven into clipping, then the total of all distortions appearing in the audio range (20–20kHz) due to transient or repetitive waveforms with frequency components inside or outside the audio range will be at least 80dB below full rated power. If the major portion of the input energy is wanted signal then −80dB (0·01%) represents the maximum possible distortion on programme.

Signal Input Overload
Instantaneous recovery up to +20dB overload.



Crosstalk
(input loaded by 1kΩ)
80dB at 100Hz.
70dB at 1kHz.

Hum and Noise
(input loaded 1kΩ)
'A' weighted −95dB ref full power.
Unweighted (15·7kHz measurement bandwidth) —90dB ref full power.

Protection
The amplifier is suitable for use under the most arduous music conditions and is electrically protected by current limiters; 7 amps in phase current at peak voltage and 3½ amps at zero voltage. Shorting both outputs simultaneously on signal for an extended period (minutes) is not protected.

Stability
Unconditionally stable with any load and any signal.

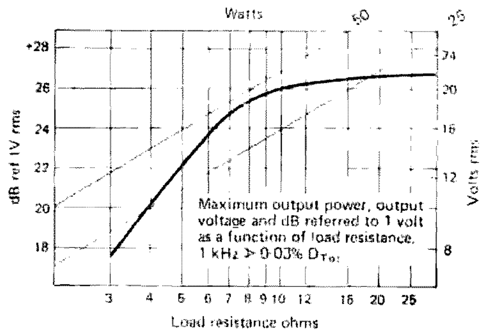
Power Input
110–130V or 220–240V, 50–60Hz, 30–350 watts depending on signal level.

Weight
9Kg.

Dimensions
Width 340·5mm
Height 115mm
Depth 195mm (plus 38mm for connectors)

Quad 303 specification
Measurements apply to either channel, with or without the other channel operating.

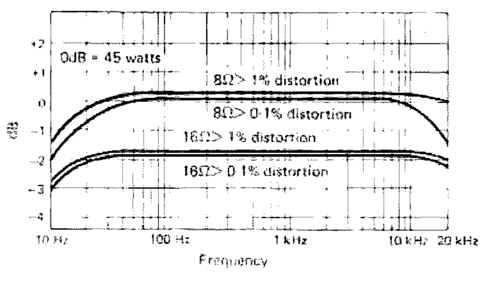
Power Output
The amplifier is intended for use with loudspeakers of 4–16Ω nominal impedance.



Power output and distortion for various frequencies. Continuous sine waves into 8Ω resistive load.
100Hz any level up to 45 watts
<0·03% Dtot
1kHz any level up to 45 watts
<0·03% Dtot
10kHz any level up to 45 watts
<0·1% Dtot

For other impedances and frequencies see graphs.

Note: In addition to the performance into a resistive load R, the amplifier will maintain full voltage within the same distortion rating irrespective of the phase angle of the load.



Output Internal Impedance
0·3Ω in series with 2000μF and 6μH.

Frequency Response
Ref. 1kHz.
−1dB at 30Hz and 35kHz into 8Ω.

Signal Input Level
0·5V rms ± 0·5dB for 45 watts into 8Ω.
Amplifier loads the input by 22kΩ in parallel with 60pF.

Signal Input Overload
Instantaneous recovery up to +20dB.

Crosstalk
(Input loaded by 1kΩ). 30Hz 10kHz > 60dB.

Hum and Noise
(Input loaded by 1kΩ).
'A' weighted −100dB ref full power.
Unweighted (15·7kHz measurement bandwidth) −95dB ref full power.

Protection
This amplifier is suitable for use under the most arduous music conditions with speakers of impedances between 4 and 25Ω, and for high level sine wave duty with loads of 8Ω or greater.

Stability
Unconditionally stable with any load.

Power Input
100–125 or 200–250V. 50–60Hz, 40–200 watts depending on signal level.

Weight
8·2kg.

Dimensions
Width 120mm
Height 159mm
Depth 423mm (plus 38mm for connectors).

Quad II Power Amplifier Specification

Figures for response, distortion, sensitivity and background are the pass figures on final test.

Power Output
15 watts throughout the range 20–20,000 c/s.

Frequency Response
Within 0·2dB 20 20,000 c/s.
Within 0·5dB 10–50,000 c/s.

Distortion (measured at 12 watts output)
Total 3rd and higher order: less than 0·1% at 700 c/s.
Higher order alone: less than 0·03% at 700 c/s.
Valve mismatching up to 25% (introducing 2nd harmonic) not to cause distortion to exceed 0·18%.
Total distortion at 50 Hz not to exceed 0·25%.

Input
Sensitivity: 1.4V rms for 15 watts output.
Load imposed on input: 1·5 Megohms in parallel with 10 F.

Background
80 dB referred to 15 watts.

Output Impedances (15 ohm and 7 ohm)
Effective output resistance: 1·5 ohm for 15 ohm output.

Power Supplies
INPUT:
200–250V AC single phase (or 95–125V AC 40–80 c/s.
90 watts consumption (excluding control unit-tuners, etc.)
HT and LT supplies available for external equipment:
330V 40mA.
6·3V 4A (heater C.T. to chasis).
VALVES:
2 × EF.86 (Z.729 or 6267), 2 × KT.66, 1 × GZ.32 (54KU or 5V4G).

PAIRED COMPARISON TEST RESULTS

Comparison	Quad II/405			Quad II/303			Quad 303/405			Same Amplifier	
	Prefer II	Prefer 405	No Preference	Prefer II	Prefer 303	No Preference	Prefer 303	Prefer 405	No Preference	Preference	No Preference
Listener a	5	4	15	7	6	11	5	6	13	11	13
Listener b	2	2	20	1	3	20	4	3	17	3	21
Listener c	3	6	15	5	3	16	4	1	19	7	17
Listener d	4	9	11	2	4	18	7	4	13	8	16
Listener e	2	3	19	2	2	20	3	1	20	3	21
Listener f	8	7	9	8	10	6	7	4	13	14	10
Group results	24	31	89	25	28	91	30	19	95	46	98

When statistically analysed using the 50% Probability Test none of these results indicate either on a group basis, or an individual basis, that there are any audible differences in the performance of the three amplifiers.

The Quad comparative amplifier tests

Tuesday, 21st March 1978

Conclusions

Three different power amplifiers—Quad II, 303 and 405 were subjectively compared for sound quality, particular regard being paid to:

1 All amplifiers were operated below overload.

2 Proper care was taken to ensure that any differences in the sound comparisons were *solely* attributable to the amplifiers.

3 A sufficient number of judgements were made to satisfy proper statistical analysis.
No differences or preferences were detected.

Acoustical Mfg. Co. Ltd.,
HUNTINGDON.

Printed by Heffers Printers Ltd Cambridge England