

ON THE ANALYSIS OF GRAVITATIONAL WAVE DATA

S. P. Lawrence
E. Nahmad-Achar*
B. F. Schutz

Department of Applied Mathematics & Astronomy
University College, Cardiff
P.O. Box 78, Cardiff, U.K.

ABSTRACT. An analysis of data was carried out from the interferometric gravitational wave detector prototype at Glasgow, to which artificial signals analogous to those emitted by coalescing binary systems were applied. We report on the ability to detect those signals, and the methods used, as well as on a full account of the detector's output noise analysis.

1. INTRODUCTION

The possibility of building detectors capable of responding to gravitational waves brings of necessity the development and improvement of techniques to analyse their output. In this paper, we report on the analysis of data from the interferometric gravitational wave detector prototype at Glasgow. The idea was to see if we could detect signals associated with coalescing binaries, that were artificially added to the detector's output by the Glasgow group. Not only do we answer the above affirmatively, but we also give a full account of the output's noise analysis.

Section 2 below describes the contents of the data sent to us by the Glasgow group. Section 3 describes the Fourier analysis and compares the results obtained from the Glasgow data to those expected theoretically. The noise analysis is described in Section 4, and we conclude this paper with a few comments.

2. DESCRIPTION OF THE DATA

The output data from the Glasgow prototype detector consists of two independent streams, each associated with one of the two arms of the

* Present address: Centro de Estudios Nucleares, UNAM, Ciudad Universitaria, Apartado Postal 70-543, 04510 Mexico, D.F., Mexico.

detector. (For the technical details of how the data is taken, cf. the paper by Norman L. Mackenzie in this volume.) One of the streams, called the *primary data*, is used as housekeeping data, and it contains the evidence of drop-outs (i.e. when the detector is not locked properly; data obtained in these intervals is to be rejected), the time records, and a large-amplitude version of the artificially applied chirps. The other stream, called the *secondary data*, contains the detector's output itself (which is essentially noise in this case), the artificially applied signals corresponding to chirps from coalescing binary systems (see below), and calibration signals which consist of equal-amplitude pulses applied at different frequencies. These calibration "combs" are used to compensate for the fact that the secondary data are related (for instrumental reasons) to the detector's response in a way that is not uniform throughout the frequency spectrum.

The signal in both streams was sampled at 6000 Hz for 25 minutes (to avoid aliasing, the data itself was filtered above 3000 Hz before sampling), and was stored as 1-byte data. When searching for signals in noise one byte data is ideal. However its use does have some restrictions on the analysis of the amplitude statistics of the noise because of saturation or truncation effects.

The artificial signal applied to the detector's output, corresponding to that which one would obtain from a coalescing binary system, is a chirp which increases in amplitude and frequency with time in a well-defined way. If we place the chirp in a standard form in which it has unit amplitude and frequency 100 Hz at $t = 0$, then we have

$$\text{chirp}(t) = [1 - 0.34\rho t]^{-\frac{3}{2}} \cos \left\{ \frac{320\pi [1 - (1 - 0.34\rho t)^{\frac{3}{2}}]}{0.34\rho} + \phi \right\} \quad (1)$$

where ρ is the *mass parameter*, which is given by $\rho = (\text{reduced mass}) / (\text{total mass})^{2/3}$ (in units of solar masses) and characterises the chirp, t is measured in seconds, and ϕ is the phase of the wave when it reaches 100 Hz. If we confine the chirp to the interval (t_0, t_1) during which its frequency rises from f_0 to f_1 , then we have

$$t_0 = \frac{[1 - (f_0/100)^{-\theta/3}]}{0.34\rho}, \quad t_1 = \frac{[1 - (f_1/100)^{-\theta/3}]}{0.34\rho},$$

Figure 1a shows a chirp with $\rho = 10$ in a frequency range 300–2000 Hz (this value for ρ was only chosen to make the graph clearer), while Figure 1b shows the real part of its Fourier transform. Figure 1c shows the power spectrum of a chirp with $\rho = 0.15$; again, this value was chosen for clarity.

The signals actually applied to the detector were chirps in the range 300–2100 Hz. All the applied chirps had the same mass parameter, which was unknown to us, as was their location in the data streams.

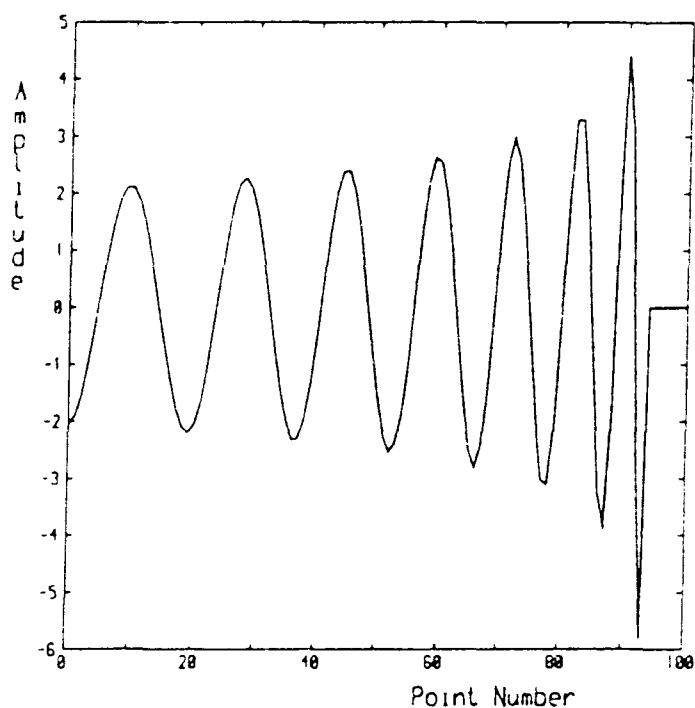


Figure 1a. Theoretical chirp for $\varphi = 10$. This large value for φ gives such a short chirp ($1/60$ s) that the sharp increase in frequency towards the end is not noticeable.

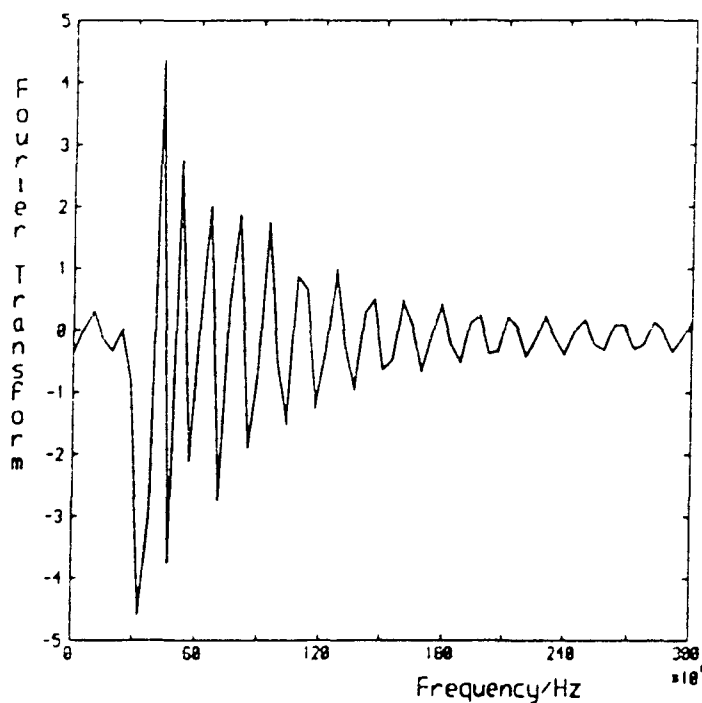


Figure 1b. Real part of the Fourier transform of the chirp in Fig. 1a.

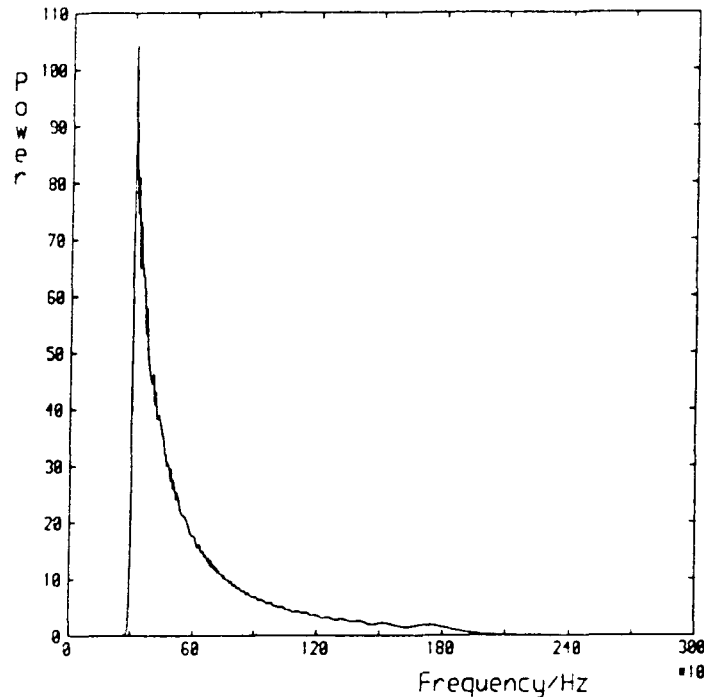


Figure 1c. Power spectrum of a chirp with $\rho = 0.15$.

3. FOURIER ANALYSIS

In order to detect the chirps embedded in the data, we built a family of theoretical chirps obeying Eq.(1) and starting from 300 Hz, with different mass parameters. (We chose 300 Hz as the lower limit because the Glasgow detector has relatively poor sensitivity below this frequency.) These were then used as templates with which to correlate the detector output. For the explicit calculation of the correlations we used the fast Fourier transform (FFT) algorithm: we know that, if $F[g]$ denotes the Fourier transform of the function g , then

$$F[\text{corr}(g,h)] = F[g]F^*[h] \quad (2)$$

where an asterisk denotes complex conjugation. The direct correlation of N points takes of order N^2 operations, while the FFT of N points takes of order $N \log_2 N$. [We found that the FFT of a set of 4096 points take 0.87 s on a MicroVax II; writing the result to a (direct access) file takes a further 0.20 s, and reading the values from a file takes 0.01 s, making a total of 1.08 s. An improvement of speed by a factor of about 100 is needed if the output of the planned detectors is to be analysed in real time.]

Using these methods we were able to detect the chirps embedded in the secondary data. After finding that the mass parameter lay in the range $0 < \rho \leq 0.1$, we wished to fix ρ more accurately. To do this, a small subset of mass parameters in the above range was chosen. The correlations for these mass parameters were calculated and the value of ρ which yielded the maximum correlation was then used, as the centre of a smaller interval. The range was continually narrowed in this way,

until it was felt that no advantage could be gained by having further accuracy; i.e. that any change in the maximum correlation would be masked by fluctuations due to the noise. In all, ten filters were needed.

The best fit was obtained with $\rho = 0.087 \pm 0.001$. Discussions with the Glasgow group later revealed that 0.087 was exactly the mass parameter that they had used in constructing their artificial chirps. This initial 'blind' test of the correlation method was therefore successful. (We should note that, because of the very low value of this mass parameter, resulting in a very long chirp, only that portion of the template between 300 and 400 Hz was used for the final correlations. This allowed for more manageable subsets of data.)

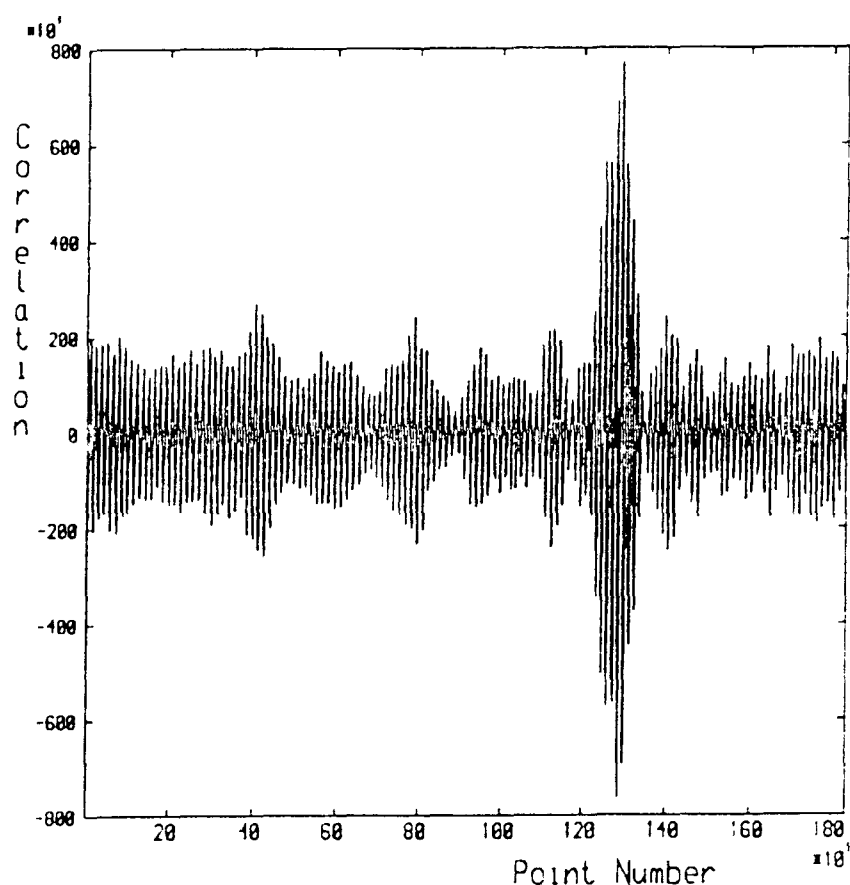


Figure 2. Correlation of theoretical chirp of $\rho = 0.087$, with subset of secondary data.

Figure 2 shows the correlation (not normalised) between the theoretical chirp with $\rho = 0.087$ and a section of the secondary data containing an applied chirp. The plot is made against record point number, and the maximum peak corresponds to the region in the embedded chirp where its frequency is approximately 400 Hz.

Figure 3 shows the log of the power spectrum of the record in which the maximum correlation in Figure 2 occurs. One notes a peak at approximately 380 Hz, and this may be due to the presence of an applied chirp. However, since corresponding peaks occur in other records at similar frequencies (cf. Figure 5), we cannot easily judge whether a chirp is present in Figure 3, or whether such peaks are due to instrumental bias. For this reason, correlations are to be preferred to power spectral methods in detecting chirps in our data.

Finally, one notes that there are equally spaced, narrow peaks in Figure 3, corresponding to interference from the mains supply at 50 Hz and higher harmonics.

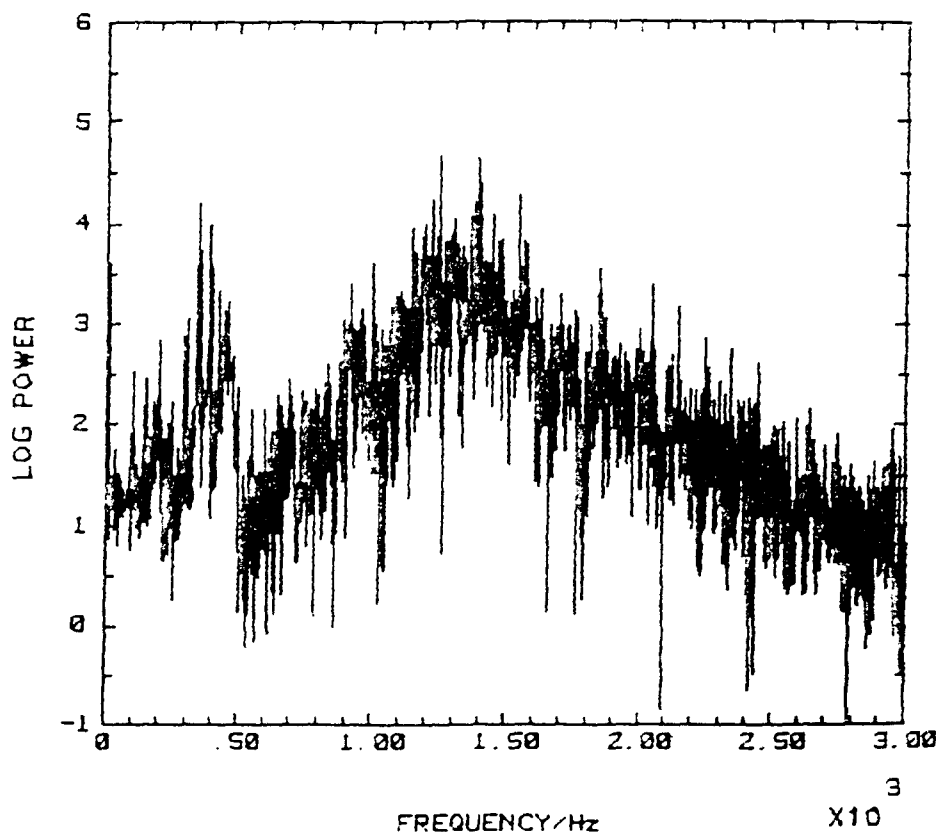


Figure 3. Log of the power spectrum of a record in which an applied chirp occurs

4. NOISE ANALYSIS

To examine the performance of the detector, we studied the output data itself, which essentially consisted of noise. Ideally this noise would be Gaussian, with zero mean. However instrumental effects will always

cause small zero offsets, and other non-Gaussian noise sources may be present. In fact, taking the whole data set (excluding calibration combs and drop-outs) we obtained

$$\text{mean} = -11.79$$

(3)

$$\text{standard deviation} = \sigma = 26.16.$$

This non-zero mean introduces a zero frequency, large amplitude peak in the power spectra of all records of the secondary data, (Figure 3 has been re-scaled so this peak is not present there). This mean has been subtracted from the data before doing the analysis discussed in this section, to account for this problem.

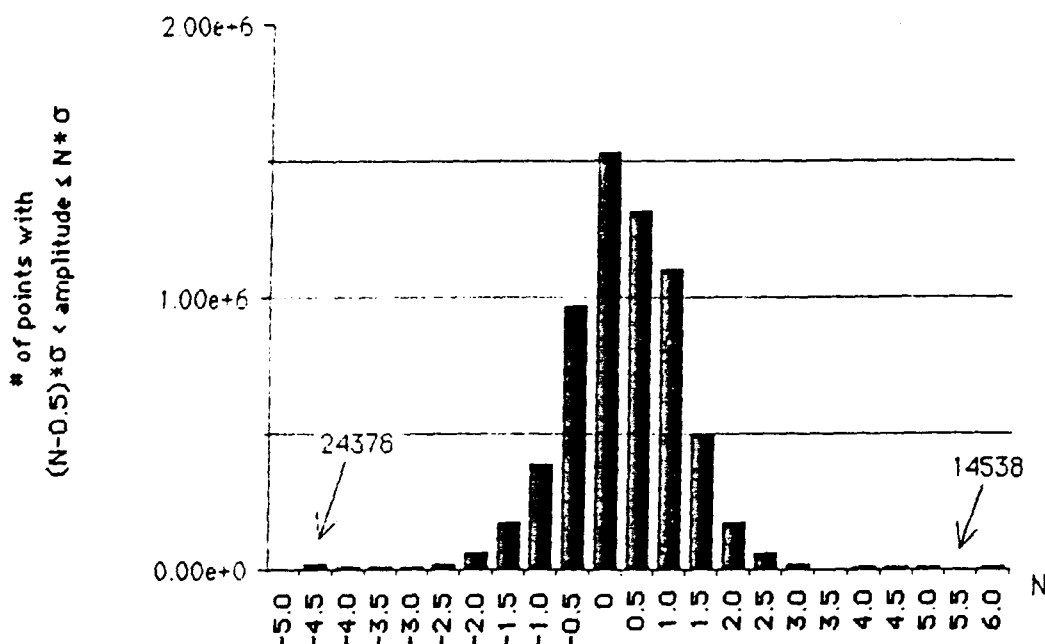


Figure 4. Histogram of time data.

Figure 4 shows a histogram of the centralised (i.e. mean-subtracted) time series. The number of points having amplitude greater than N times σ , but less than $(N+0.5)$ times σ , is plotted against N . Note that, while the central region of the histogram looks roughly Gaussian, two peaks show a relatively large number of points at -4.5σ and 6σ . This is most likely due to the limiting of the dynamic range of the data (cf. Section 2), by which all amplitudes are truncated at the absolute value of 127.

We have also mentioned interference at 50 Hz and higher harmonics from the mains supply. The power spectrum of an arbitrary subset of the

data not only shows this interference, but also the non-uniform instrumental response, peaking near 1200 Hz (see Figure 5). This bias can be compensated for by using the calibration signals (combs) applied to the data (cf. Section 2).

However, the data also contains non-white noise and to allow for this one could use a "local estimate" of σ : for every small subset of data (approximately 100 points), calculate its σ value and look for peaks inside the subset, relative to this σ . An analysis along these lines should not only allow a study of the statistics of the power spectrum of the raw data, but would also provide a way to set an upper limit on the detector's sensitivity. Such analysis is currently being carried out.

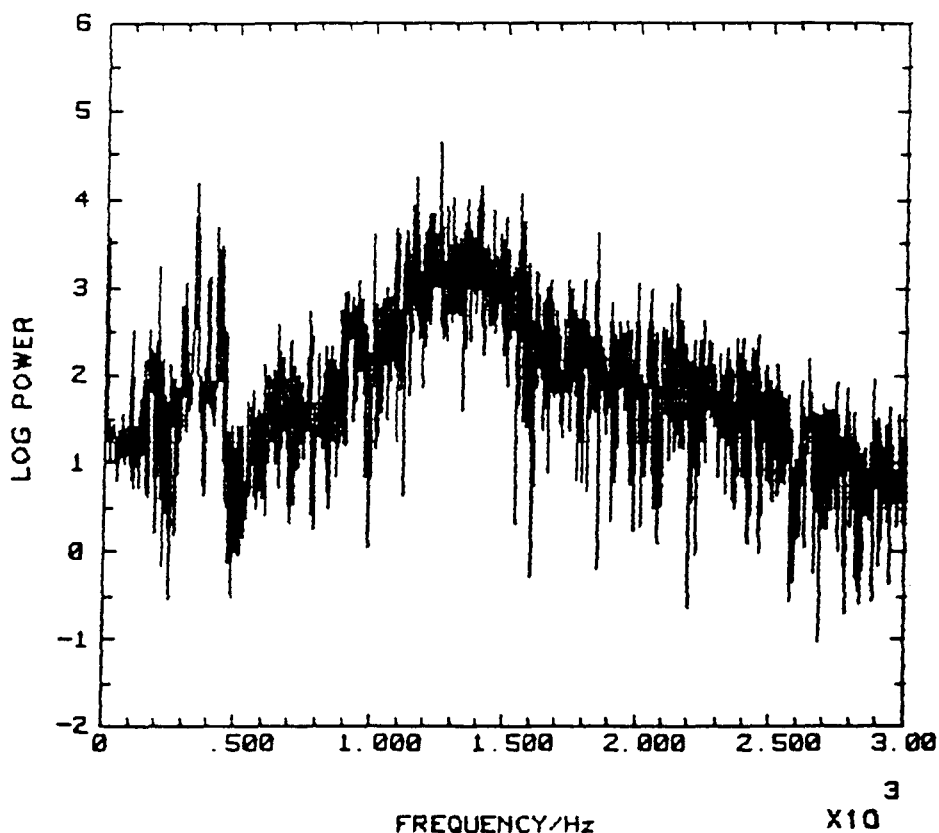


Figure 5. Log of the power spectrum of arbitrary record of time series in which a chirp is not present.

5. COMMENTS

We have shown that, in a blind test of simple filtering techniques, we have been able to detect the artificially applied chirps in the Glasgow data and to determine the mass parameter of the chirp accurately. The correlation method we used can therefore be expected to be useful in the

detection of true gravitational wave signals in future detectors' output. It must be said, however, that a continuously running detector will output vast amounts of data, and since one cannot predict (at least with sufficient accuracy) the kind of event that will take place, all the data has to be correlated with more than one family of filters (for coalescing binaries, for supernovae explosions, for neutron star and black hole collapses, etc.). This asks for both real-time analysis of the output and extremely fast hardware-software combinations that can handle it.

Our experience with this experiment has shown that any attempt to develop an automatic data-analysis system for real-time analysis of long data streams will have to find ways to eliminate or cope with instrumental features in the data. In this case these included interference from the electricity supply, and truncation and centring errors in the 1-byte data, and the non-uniform frequency response of the output data. In the long run, a standard format for data storage would allow the software to run more efficiently (e.g. by choosing the record lengths to be an integer power of 2), and this together with agreement on the information to be stored in the housekeeping data, should facilitate data interchange between the different groups.

ACKNOWLEDGEMENTS

The authors would like to thank the Glasgow group for supplying the data and for valuable discussions. Computations were carried out on a DEC MicroVax II supplied under SERC grant GR/E46158.