A Reprint from the

PROCEEDINGS

Of SPIE-The International Society for Optical Engineering

Volume 396

Advances in Laser Scanning and Recording

April 19-20, 1983
Geneva, Switzerland

Maximization of recording density obtainable in Te-alloys

Kees A. Schouhamer Immink, Ronald M. Aarts
Philips Research Laboratories
5600 JA, Eindhoven, The Netherlands

Maximization of recording density obtainable in Te-alloys

Kees A. Schouhamer Immink and Ronald M. Aarts

Philips Research Laboratories
5600 JA, Eindhoven, The Netherlands

Abstract

We report on high density recording experiments of digital information in Te alloys on pregrooved discs. The recording and reading of information is done on a recorder fitted with an AlGaAs laser. We describe experiments with modulation systems using pit-length modulation based on runlength-limited codes. Runlength-limited sequences were adopted as a modulator output because of the fact that this class of restricted sequences has a great impact in magnetic and optical recording.

State-of-the-art high power solid-state lasers can emit a light pulse of sufficient energy for only a limited time and can therefore only be used in a pulsed mode. Pit-length modulation is achieved by adjusting the rotational velocity of the disc and laser pulse write frequency in such a way that oblong pits of overlapping monoholes result.

We demonstrate the feasibility of recording densities of up to 1 Mbit/mm^2 with the application of pit-length modulation schemes.

Introduction

In this paper we consider the feasibility of high density optical recording in Te alloys. The recording and reading of information is done on a recorder equipped with an AlGaAs laser. All our experiments are done on pregrooved discs with a track pitch of 1.7 μm.

We describe experiments with pit-length modulation. All the modulation systems we considered are based on runlength-limited sequences. Runlength-limited sequences were adopted as a modulator output because of the fact that this class of restricted sequences has a great impact in magnetic and optical (read-only) recording as the Compact Disc Digital Audio System.[1,2,3,4]

In Section 2 we give a brief definition and theory of runlength-limited sequences. Section 3 gives theoretical results of the maximum achievable information density and the sensitivity to parameter tolerances. A description of the experiments is given in Section 4.

2. Definition of runlength-limited binary sequences

The theory of binary sequences with restrictions on minimum and maximum runlength goes back to Kautz[5], Tang and Bahl[6]. For an exhaustive treatment of this subject the reader is referred to ref. 6. Their most important results are summarized here.
We adopt Tang's definitions:
A dk-limited sequence simultaneously satisfies the following conditions:

a. d-constraint - two logical ones are separated by a run of consecutive logical zeros of at least d.

b. k-constraint - the length of any run of consecutive logical zeros is at most k.

A sequence satisfying the d- and k-constraint is called a dk-sequence. Sequences only satisfying the d-constraint are called d-sequences. We derive a runlength-limited binary sequence with at least (d+1) consecutive zeros or ones and at most (k+1) consecutive zeros or ones by integrating modulo 2 a dk-limited sequence. In this way the "ones" of a dk-sequence indicate the position of a transition zero to one or one to zero of a runlength-limited sequence. In ref. 6 recursion equations are derived for the number of distinct dk-limited sequences of block length n as a function of d and k. If for convenience we restrict ourselves to d-limited sequences, the number of distinct binary sequences N of block length n is given by the following equation.

$$N(n) = \begin{cases} n+1 & 1 \leq n \leq d+1 \\ N(n-1) + N(n-d-1) & n > d+1 \end{cases} \qquad (2.1)$$

The asymptotic information rate R of a dk-sequence is determined by the specified constraints and is given by

$$\lim_{n \to \infty} \frac{1}{n} \, {}^{2}\log N(n) \qquad (2.2$$

For large n the number of distinct d-sequences N(n) behaves as $\alpha\lambda^n$, with λ given by the largest real root of

$$z^{d+1} - z^{d} - = 0 \qquad (2.3)$$

The maximum information rate (in short rate) R is then simply given by

$$R = {}^{2}\log \lambda \qquad (2.4$$

Similar relations can be derived for sequences with a d and/or k constraint. The process of modulation maps the input data stream onto the runlength-limited output stream. In general m consecutive databits are mapped onto n consecutive channelbits. The existence of a maximum asymptotic information rate R merely states that $m/n < R = R(d,k)$ for any finite m and n. If one wishes to transmit a certain fixed amount of information per time unit over a dk-limited noiseless channel, then the channel clock should run at least 1/R times faster than the data clock to compensate for d- and k-constraint. In other words, a channel bit takes a time which is shorter by a factor of at least R than that needed for a data bit. The minimum physical distance per data bit of the runlength-limited sequence generated by modulo 2 integrating a dk-sequence is now given by

$$T_{min} = (d+1)R(d,k) \qquad (2.5)$$

In table I we have listed for some values of d the minimum distance T_{min} between transitions and the rate R. Note that d=0 is the uncoded (NRZ) case. Both R and T_{min} are specified per data bit length.

d	R	$T_{min} = (d+1).R$
0	1.0	1.00
1	0.69	1.38
2	0.55	1.65
3	0.46	1.84
4	0.41	2.05
5	0.36	2.16

Table I. Rate R and T_{min} versus d for fixed data rate.

The values of R and T_{min} in table I are theoretical maxima. In this paper we will not discuss an actual implementation of encoders and decoders. Good algorithms can be found in[6,7,8] and it suffices to state that practical modulators with finite dimensions reach the maxima of table I to within 90%.[7]

We note from the table that application of dk-sequences enables us to increase the minimum time between transitions T_{min}. The rate R (or channel bit time per data bit time) decreases with d. Qualitatively we may state that T_{min} is related to the highest frequency of the modulator and hence to the maximum attainable information density. Table I shows clearly the possible trade-off between T_{min} (related to the highest frequency in the runlength-limited sequence) and the timing accuracy (timing window) RT.

To get some idea of the trade-off, Fig. 1 depicts the eye-patterns for RLL codes with d=0, 1, 2 and 3. In this figure the minimum time between transitions of the modulation stream T_m is now fixed. Consequently the data rate i.e. the number of data bits transmitted per unit time is different. From eq. 2.5 and table I we derive that, if the d=0 code transmits 1 data bit/s, then the codes with d=1, 2 and 3 transmit 1.38, 1.65 and 1.84 data bit/s, respectively. In the figure we note clearly the decreasing eye-opening in both time and amplitude with increasing d. In the next chapter we shall study quantitatively the spectral and bandwidth properties of the runlength-limited sequences.

3. Maximization of information density

3.1. Bandwidth properties of runlength-limited sequences

We assume that the runlength-limited sequence is transmitted over a linear, bandwidth-limited channel with impulse response h(t). Furthermore we assume that the channel output is corrupted with additive Gaussian, zero mean, noise n(t). Hence the output is the sum of the convolution of the channel impulse response with the runlength-limited sequence and the noise

$$r(t) = \int_{-\infty}^{\infty} h(t')s(t-t')dt' + n(t) \qquad (3.1)$$

where

r(t) is the detected output signal
n(t) is noise
s(t) is the RLL-code input signal
h(t) is the impulse response of the channel

In this paper we restrict ourselves to simple, non-equalized amplitude detectors, so that it is only the total noise power that is important and not the exact shape of the spectrum.

With this linear channel model we can calculate the channel bit error (BER). A conceptually simple algorithm was presented by Tufts and Aaron[9]. They assumed the intersymbol interference to be confined to m data bits or equivalently n = m/R channel bits. In other words, we truncate the impulse response and ignore the contribution of the data m/(2R) bits away from the centre of the impulse. Corresponding to the N(m/R) legal dk-sequences of length m/R, there are at most N(m/R) distinct "eye-openings" at the sampling moments. The conditional error probabilities are computed for each of the run-length-limited sequences and then averaged with respect to the probability of occurrence of these sequences. This procedure has a wide field of application: any shape of the impulse response, negative or positive sidelobs can be treated thus. One major disadvantage of this procedure is the exponential growth of the computational effort with N(m/R). A computationally more straightforward method is derived on the basis of the observation that the bit error rate is dominated by the smallest, worst-case eye opening. A good approximation of the bit error rate is possible by calculating the worst-case eye opening and then calculating the error probability based on this eye opening only. If we ignore baseline wandering due to AC-coupling of the modulation stream, then the worst-case eye opening is the repetition of the minimum and maximum runlengths T_{min} and T_{max} respectively (See Fig. 2). This approximation is only valid if h(t) is strictly positive. The channel output with the worst-case input pattern can now be easily calculated with eq. (3.1), assuming the amplitude of the RLL sequence to be unity. Note that the eye-opening is the output signal r(t) sampled at the centre of the channel bit of length RT.

$$\text{eye} = \int_{-\infty}^{-RT/2} h(t)dt - \int_{-RT/2}^{(T_{min}-R/2)T} h(t)dt + \int_{(T_{min}-R/2)T}^{\infty} h(t)dt \qquad (3.2)$$

Assuming a symmetric impulse response or h(t) = h(-t), eq. (3.2) simplifies to eq. (3.3)

$$\text{eye} = 2\int_{0}^{RT/2} h(t)dt - 2\int_{(T_{min}-R/2)T}^{\infty} h(t)dt$$

For optical recording many approximations are possible to obtain a useful model of the impulse response of the read-out mechanism.[10] We base our discussions on a Gaussian roll-off frequency characteristic. Such a model is also often adopted in magnetic recording.[11,12] We ignore any radial contribution, so we adopt a completely linear one-dimensional read-out model.

We first assume that the impulse response is given by

$$h(t) = \frac{1}{\sqrt{\pi}} B\exp(-B^2t^2) \qquad (3.4)$$

The cut-off frequency B is given by

$$B \cong 2.8\ NA * v/\lambda \qquad (3.5)$$

where

v = linear velocity of the disc
NA= numerical aperture of the objective lens
λ = wavelength

We now find, using eq 3.1

$$\text{eye} = \text{erf}\left(\frac{R}{2\sigma}\right) + \text{erf}\left(\frac{R}{\sigma}\left(d+\frac{1}{2}\right)\right) - 1 \qquad (3.6)$$

where

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x \exp(-z^2)dz; \quad \text{erf}(\infty) = 1$$

and σ = normalized information density given by $\sigma = /BT$
T = data bit length.

In Fig. 3 we give the minimum worst-case eye-opening as a function of the normalized information density for d and consequently for T_{min} as a parameter. All these calculations are based on the theoretical maximum rate R and T_{min} of the theoretical runlength-limited codes (see table I). We notice from this figure that the eye-opening for a certain d decreases with the information denisty σ . For a value d' > d we notice that the eye-opening is initially smaller at low densities, but then decreases more slowly. The result is that at a certain density both eye-openings are equal and for still larger densities the eye-opening is larger for d' > d. It is quite clear from this figure that if we wish to design a high-density modulation system, then the choice of a certain d and hence T_{min} depends on both the minimum tolerable eye-opening and the information density. We note that it is only worthwhile to consider a d=2 over a d=1 system if we may tolerate a minimum relative eye height smaller than 28 percent (-11 dB).
For the Gaussian noise model the channel bit error rate (BER) can be approximated by

$$BER \approx \tfrac{1}{2} \left\{ 1 - \mathrm{erf}\left(\frac{A.eye}{\sqrt{2}N}\right)\right\} \qquad (3.7)$$

where A = amplitude of the runlength-constrained signal
and N^2 = variance or power of the additive noise.

The signal-to-noise ratio is simply given by: A^2/N^2 (the runlength-limited sequence is binary-valued with amplitude A). The actual data bit rate (due to error propagation) is a function of the channel bit error rate. This function depends on the typical implementation of the mapping of data onto channel bits and vice versa.

With eq. 3.7 and a given channel bit error rate, we can calculate the minimum tolerable eye-opening as a function of the signal-to-noise ratio. Most certainly other factors than noise are operative, such as tolerances in the transmission path, governing the minimum tolerable eye-opening. For a given minimum tolerable eye-opening we now proceed to calculat the maximum information density as a function of d and T_{min}. In Fig. 3 we draw horizontal lines at the specified tolerable eye-opening and determine the intersections with the other curves. This results in the maximum achievable information density versus d (or T_{min}) for some specified signal-to-noise-ratio (See Fig. 4). The specified bit error rate is 10^{-3}.

In Figure 4 we notice maxima in information density for some value of d. For small d values a high intersymbol interference occurs which limits the information density. At large d values, the high accuracy of the transition positioning also sets a limit to the maximum density. We note that for an SNR = 26 dB a flat maximum in information density occurs for d = 2 and 3. Note that the difference in maximum attainable information density for a given tolerable eye height is small. In principle we can infinitely increase the information density for the given Gaussian impulse response, if the signal-to-noise ratio is correspondingly improved. If for practical reasons a minimum eye-opening of say -20 dB is a limit (tolerances, etc.), then systems with d > 5 should not be considered.

4. Description of the experiments

4.1. The disc

The recording of information is done by locally heating the Te-alloy layer. Figure 5 shows a cross-section of the actual disc. A 2P (photo-polymer) layer with a groove structure (1.7 μm track pitch) in its top surface has been deposited on a glass substrate. The Te-alloy layer has been flash evaporated on the 2P layer. The entire disc can be sandwiched with a second glass or plastic substrate for protection of the sensitive layer (see for example.[13,14]

4.2. The optical recorder[13]

The optical recorder[13] has to perform two functions:

a) local heating of the bit locations
b) detection of the pits created

In Figure 6 a schematic drawing of the optics is shown. The light source is a high-power AlGaAs laser with a wavelength of 840 nm. Approximately 40% of the light output is usefully captured by an objective with a numerical aperture of 0.3. The parallel beam traverses a polarizing beamsplitter and a $\lambda/4$ plate and is focussed onto the disc by an objective with N.A.= 0.6. The half-width of the lightspot is slightly smaller than one micron. When recording information, the laser is driven by a pulse of 50 nsec duration at intervals of at minimum 250 nsec. The peak power is 60 mW; owing to losses in the light path, 10 mW is available in the focussed light spot.
A pit generated by a single light pulse has a circular shape with a diameter of typically 1 micron. Oblong pits are generated by applying several light pulses at 250 nsec. intervals.

Figure 7 shows an SEM microscope photograph of written pits in the pregrooved spiral on a disc. The sequence represents a typical digital signal with oblong pits of discrete lengths. The minimum pit length (not the bit length) is 1 micron and the spacing between tracks is 1.7 microns.
When detecting written pits, the laser is pulsed at a high frequency (15 MHz) with a duty cycle of 0.15 and a peak power of 8 mW. A quasi-continuous power of 0.3 mW is available on the disc and this is sufficiently low for the recorded pits not to be perturbed. Through

the polarizing beamsplitter and a semi transparent mirror the light is thrown onto a photo diode detector, see Fig. 6. After that the photo current is electronically processed. During recording and reading a small part of the light (10%) is coupled out to the tracking (push-pull) and focussing optics (Foucault double-wedge method). An automatic gain control is incorporated in order to compensate for the large difference in light power during recording and reading.

4.3. Pit length modulation

The design and building of real-time modulators and demodulators is a time-consuming activity. In earlier experiments we noticed that a maximal density study can be done quite well on the so-called "channel level" i.e. we do not have (de)modulator rules, but compare in BER measurements the channel bits. We used as a data source the "produlator", a PROM, filled with the desired bit stream, which is periodically read out. We filled 6 PROMs each with a frame sync pattern (27 bits) and a dk-sequence (561 bits) increasing from d=1 to d=6, k=10 fixed. The channel bit rate was 4 Mbit/sec. during all experiments.

4.4. Experiments

Pulse length modulation differs physically from monohole modulation mainly in two aspects:
- the present state-of-the-art solid-state laser is not capable of emitting light of the required (high) constant intensity level during the appropriate time;
- the behaviour of the sensitive layer during a relatively long exposure to high intensity.

To eliminate possible effects due to the above mentioned problems we have written oblong pits by writing overlapping monoholes. Figure 8a depicts the desired pit-form and Fig. 8d shows the pulses fed to the laser.
The length of the pit is

$$P_i = (i-1\ \ v.T_c+\delta \qquad (4.1$$

where

P_i = pit length
i = number of overlapping monoholes
v = tangential disc velocity
T_c = time between 2 successive laser pulses
δ = monohole diameter.

To reach the maximal density, the minimal pit size must be chosen equal to δ. According to Fig. 8 a minimum pit is equivalent to (d+1) pulses and hence d pulses of the sequence must be deleted. For the d = 2 sequence of Fig. 8c this leads to a pulse stream in Fig. 8d; the latter shows the pulses of the nominal duration, 50 nsec during writing, being fed to the laser. With a fixed T_c and deleting d pulses the velocity is fixed and is

$$v_{nom} = \frac{\delta}{T_c(d+1)} \qquad (4.2)$$

Deviations from the nominal velocity lead to unwanted deviations in P_i. The effect of a deviation from the nominal velocity is measured and plotted in Fig. 9.

Figure 9 is not symmetrical around $v/v_{nom} = 1$ because of the decreasing density and consequently increasing eye opening at $v/v_{nom} > 1$. The energy used for creating a hole in the Te alloy film determines its size and consequently the pit length is a function of the energy in the write pulse. The sensitivity of energy deviations in the write pulse to the BER is measured and Fig. 10 depicts the result. The figure shows a BER optimum when an energy of approximately 1 nJ during the write pulses is used. At the nominal speed and a BER of 5E-4 the maximum obtainable density was about 2 data bit/μm .

Conclusions

We have described experiments with modulation systems based on pit length modulation in Te-alloy based discs. Computer simulations show a superiority modulation system based on a d = 2 runlength-limited sequence. Experiments confirmed these results. We achieved an information density of 1 Mbit/mm^2 (1.7 μm track pitch and 0.6 μm tangential density).

Acknowledgement

The authors are indebted to Wil G Ophey, who designed the optics in the experimental set-up.

References

1. J.C. Mallinson and J.W. Miller, "Optimal Codes for digital magnetic recording", Radio and Elec. Eng., 47, p. 172-176 (1977).
2. G.V. Jacoby, "A new look-ahead code for increased density", IEEE Trans. Magn., vol. MAG-13, no. 5, pp. 1202-1204 (1977).
3. K.A. Immink, "Modulation Systems for Digital Audio Discs with Optical Read Out", Int. Conf. ASSP, Atlanta, pp. 578-590 (1981).
4. J.P.J. Heemskerk and K.A. Schouhamer Immink, "Compact Disc; system aspects and modulation", Philips Techn. Rev. 40, pp. 157-164 (1982).
5. W.H. Kautz, "Fibonacci codes for synchronization control", IEEE Trans. Inform. Theory vol. IT-11, pp. 284-292 (1965).
6. D.T. Tang and L.R. Bahl, "Block codes for a class of constrained noiseless channels", Inform. Control. vol. 17, pp. 436-461 (1970).
7. P.A. Franaszek,"Sequence-state methods for runlength limited coding", IBM J. Res. Develop., pp. 376-383 (1970).
8. G.F.M. Beenker and K.A. Schouhamer Immink, "A simple method for coding and decoding of runlength-limited sequences", IEEE Conf. on Information Theory, les Arcs (1982).
9. M.R. Aaron and T.W. Tufts, "Intersymbol Interference and Error Probability", IEEE Trans. Inform. Theory, IT-12, pp. 26-34 (1966).
10. G. Bouwhuis and J. Braat, "Video disk player optics", Applied Optics, Vol. 17, pp. 1993-2000 (1978).
11. A.S. Hoagland, "Digital Recording", Wiley (1963).
12. N.D. Mackintosh, "A superposition-based analysis of pulse-slimming techniques for digital recording", Int. Conf. on Video and Data Recording, Southampton (1979).
13. K. Bulthuis et al., "Ten billion bits on a disc", IEEE Spectrum (August 1979).
14. P. Kivits et al., "The hole formation process in Tellurium layers for optical data storage". Thin solid films, 87, pp. 215, (1982).

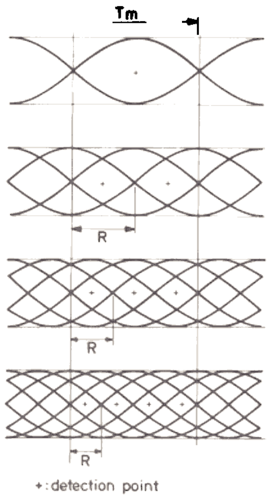

Figure 1. Eye-pattern of some runlength-limited codes for a fixed minimum time between transitions T_m.

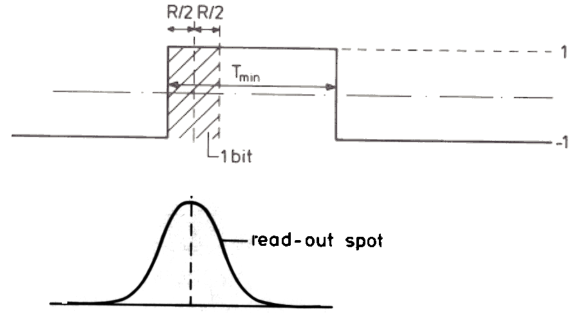

Figure 2. Worst-case input pattern and impulse response of the read-out system.

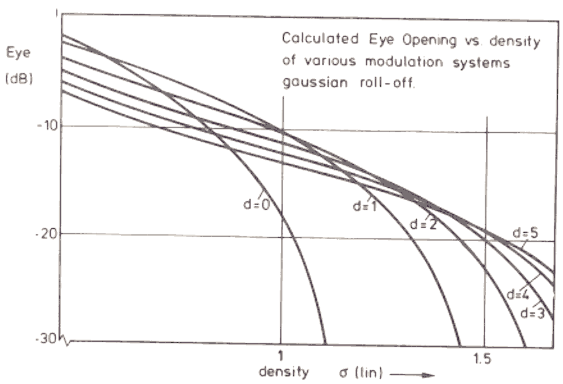

Figure 3. Worst-case eye-opening versus normalised information density with T_{min} as a parameter, Gaussian roll-off.

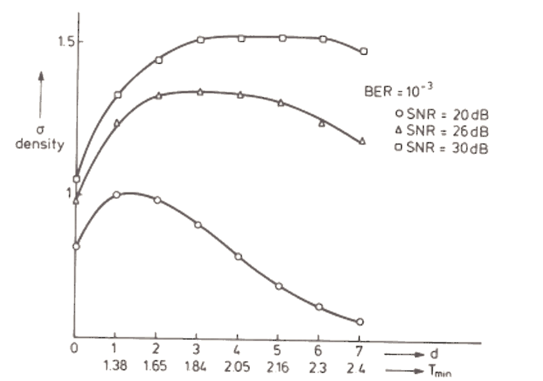

Figure 4 Maxir um ac ievable information density or a BI R of E-3 versus minimum distance betweε n tra sitions T_{min} for certain NRs, (aussi n roll-off. Substitu ing sc me pr ctical values : SNR = 26 dB, N/ = 0. , λ = 840 nm yields a maximι m den ity of about 2 data bit per ı icron

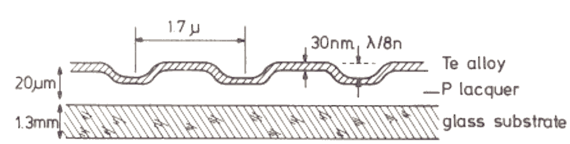

Figure 5. A cross-section of the disc.

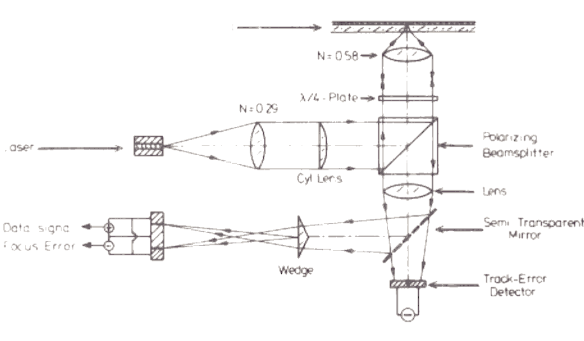

Figure 6. Schematic drawing of the optical set-up.

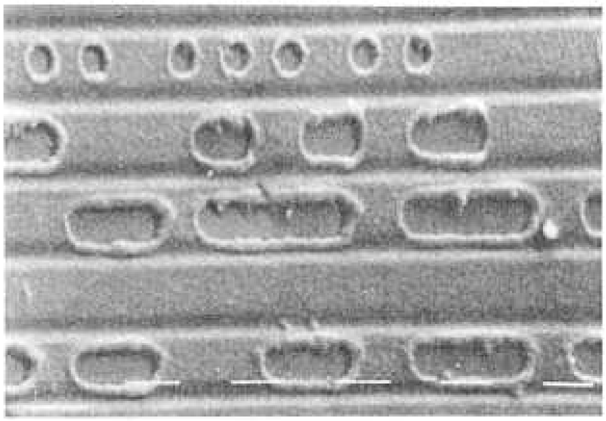

Figure 7. A scanning electron microscope photograph of a 10000 times enlarged disc sample with a Te compound layer. The upper track contains monohole information, the other tracks runlength sequences. The observation angle is 45°, with one white bar corresponding to 1 μm.

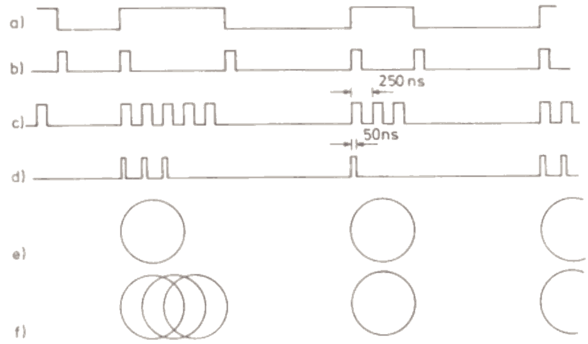

Figure 8.
a) A runlength limited sequence.
b) A dk-sequence.
c) A pulse derived from an RLL sequence used for writing overlapping monoholes.
d) Write pulses for overlapping monoholes fed to the laser (2 pulses deleted).
e) Schematic pits on the disc when monoholes are used.
f) Schematic pits on the disc when pitlength modulation (overlapping monoholes) is used.

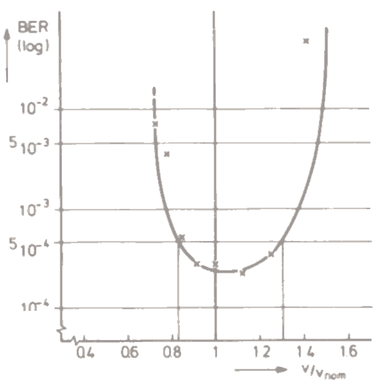

Figure 9. The BER versus velocity deviations (experimental).

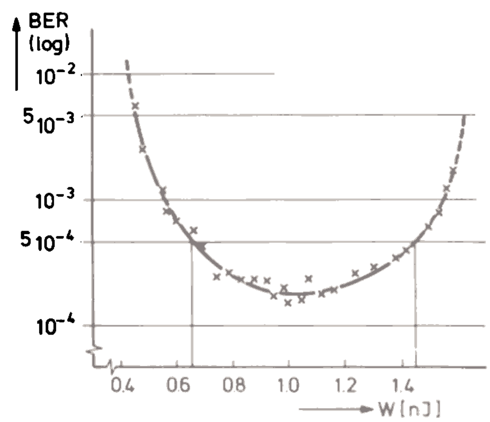

Figure 10. The BER versus energy deviations (experimental).