ISSN 1330–0016
CODEN FIZBE7

MICROEMULSIONS, L_3 PHASES AND INFORMATION LOSS IN SIGNAL REPRESENTATION[1]

STJEPAN MARČELJA

Department of Applied Mathematics, Research School of Physical Sciences and Engineering, The Australian National University, Canberra ACT 0200, Australia

Dedicated to Professor Mladen Paić on the occasion of his 90$^{\mathbf{th}}$ birthday

Received 11 May 1995

UDC 536.75

PACS 03.40.Gc

The surfaces of bicontinuous disordered microemulsion structures or L_3 phases can be described as level cuts of random Gaussian fields with a given spectrum. We consider here the difference between the entropy of these states and the known entropy of the generating random field. In another context, this difference is equal to the decrease in the Shannon information content resulting from the process of sampling and digitising of continuous stochastic signals.

1. *Level-cut random fields and microemulsion structure*

The most accurate analytical model of microemulsion structures is obtained by defining the surface dividing oil from water as a level cut of a 3-dimensional random Gaussian field [1 – 3]. For a symmetric microemulsion, the surface is defined as the locus of all points in space where the field has a value of zero (Fig. 1a). Non-symmetric phases can be described by changing the level of the cut to other values

[1]In this tribute I would like to recall the love for all aspects of physics which as beginning students we learned from Professor Paić. As I write this I am thinking of the time 35 years ago when Professor Paić lead the Zagreb University class of 1960 through the excitement of entering the world of scientific enquiry with its limitless possibilities.

(Fig. 1b). For the L_3 phase, the surface defines a membrane separating "inside" and "outside" aqueous compartments.

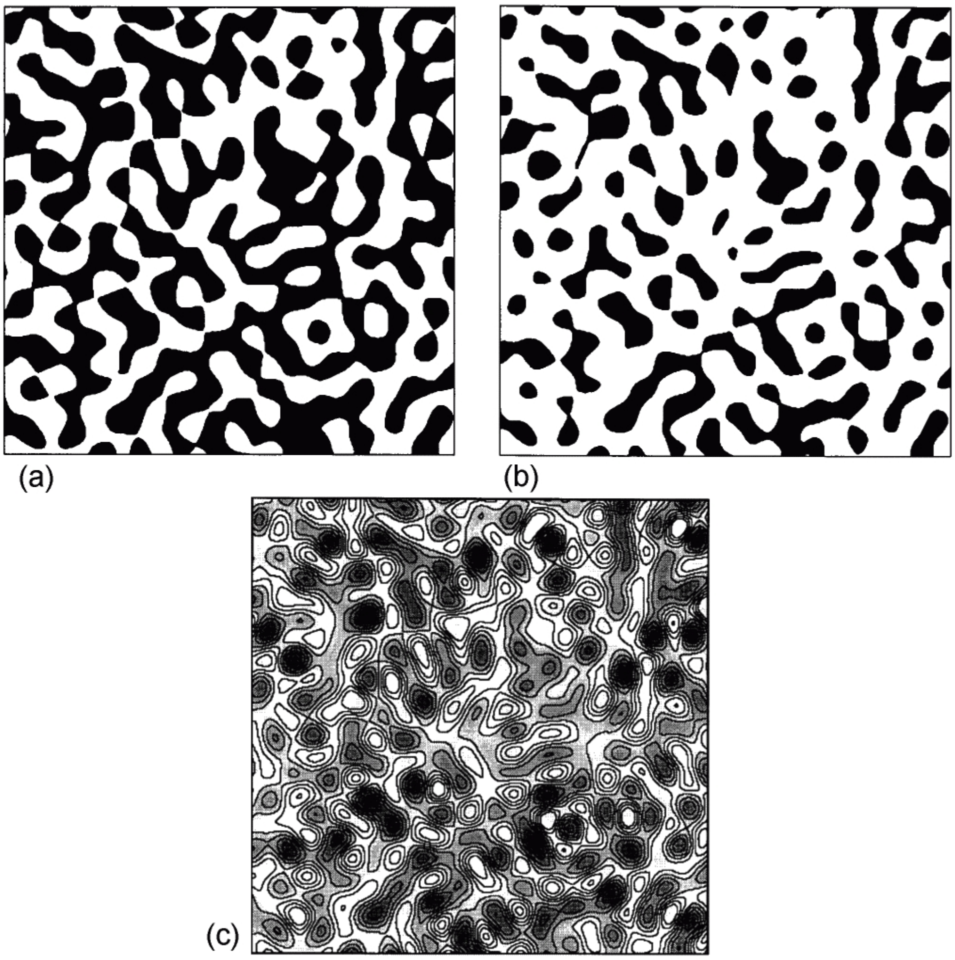

Fig. 1. Planar cross–sections of a levelled Gaussian field. The states can correspond to oil/water distribution in a symmetrical microemulsion (a); non-symmetrical microemulsion (b); or many levels of digitised signals (c).

As the generating random Gaussian field is completely described by its spectral density, average properties of the level-cut surfaces also depend only on the spectral density. The average elastic energy of surfaces was evaluated exactly using the properties of the Gaussian statistics of the generating field [2]. The entropy of surfaces is more difficult to evaluate. It was approximated by the entropy of the generating field, based on the argument that it is possible to reconstruct the 3-dimensional field from the level cut, and therefore the surface and the field contains the same information [3]. The free energy of the system was thus approximately determined and the spectrum of the field could be obtained by variational minimisation of the free energy. This completed approximate derivation of the constitutive equations of the model. For applications of the model, see Ref. 4.

The reconstruction of 3-dimensional fields from the surfaces is accurate if surfaces are defined with great precision. In numerical experiments [2], we had no difficulty reconstructing a 3-dimensional field from level-cut data using a variant of the method used to reconstruct 2-dimensional images from zero crossings [5]. But in real physical systems, surfaces are only defined with finite accuracy, and the argument relating the entropy of the generating field to the entropy of surfaces must be re-examined.

2. *Link to signal representation*

While this problem may appear to belong to a rather exotic area of chemical physics, it is in fact equivalent to a very general problem in signal processing. Continuous signals are routinely sampled and digitised. Digitising process corresponds to a cut of a continuous field at one or more levels (Fig. 1c). The accuracy of a digitised record can be a single bit – just black and white[2], a 16-bit word commonly used in recording of music, or any other desired accuracy. How much information is lost in the process?

Of course, the answer depends on the correlation structure of the original continuous signal. But for one important class of signals – the Gaussian process with a known spectrum – we expect that this general question can be answered by evaluating the entropy of the level cut Gaussian fields using the methods developed over the years in statistical physics of solids and liquids. The evaluation will provide some insight into the relationship between the correlation functions and the information content of continuous signals.

In presenting the above argument, we use the exact equivalence between the most common measure of information - the Shannon information [7] – and the thermodynamic entropy of the system. Other measures of information, like the algorithmic information of Kolmogorov and Chaitin may be more appropriate for very structured signals.

3. *Link to the Ising model*

After a level cut, the Gaussian random field becomes a spatially distributed field which can only assume two values. It is a natural expectation that such fields

[2]To increase the amount of information in the black/white image, various dithering algorithms are available (see eg. Ref. 6).

can be described by a corresponding Ising model. In a more general version, field values can assume a discrete number of states and the resulting statistics is related to the Potts model.

The expectation is borne out in the lowest order approximation, where simple results are obtained. Further investigation is needed for higher order terms and the re-summation of some classes of diagrams.

In the following, we use a discrete lattice in space, where the lattice constant corresponds to the precision with which a surface can be defined in chemical physics applications, and in signal processing applications to the sampling rate. It is assumed that the sampling rate is sufficiently high to represent the highest frequencies in the signal.

The Gaussian random field x_i is completely specified by its covariance matrix $R = (r_{ij})$. We select a single level cut at a value α and obtain surfaces as illustrated in Figs. 1a and b. The space is divided into regions of, e.g. oil and water. Probabilities of oil or water at each of n lattice sites are obtained by integrating over all field values $x \geq \alpha$ or $x < \alpha$, respectively. The probability of a particular state on the lattice is

$$p_i = \frac{(\det R)^{-\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \int\limits_{\alpha}^{\infty} \int\limits_{-\infty}^{\alpha} \dots \int\limits_{\alpha}^{\infty} \mathrm{d}x_1 \dots \mathrm{d}x_n \exp\left(-\frac{1}{2} X^T R^{-1} X\right). \tag{1}$$

$X = x_1 \dots x_n$ is a vector describing a configuration of the generating Gaussian field. After levelling, the configuration can be described with a set of discrete variables $\mu_i = \pm 1$. Using the transformation $x_i \to \mu_i x_i$ and a summation over the variables μ_i, we obtain the configurational partition function

$$Q = \frac{(\det R)^{-\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \sum_{\mu_i = \pm 1} \int\limits_{\mu_1 \alpha}^{\infty} \int\limits_{\mu_2 \alpha}^{\infty} \dots \int\limits_{\mu_n \alpha}^{\infty} \mathrm{d}x_1 \dots \mathrm{d}x_n \exp\left(-\frac{1}{2} X_\mu^T R^{-1} X_\mu\right), \tag{2}$$

where X_μ stands for the vector whose elements are $\mu_i x_i$. Integration limits in this equation can be considered as arising from Heaviside step functions inside the integrals. Expressing these functions in the integral representation gives

$$Q = \frac{1}{(2\pi \mathrm{i})^n} \frac{(\det R)^{-\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \sum_{\mu_i = \pm 1} \int\limits_C \int\limits_C \dots \int\limits_C \frac{\mathrm{d}z_1 \dots \mathrm{d}z_n}{z_1 \dots z_n} \times$$

$$\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \dots \int\limits_{-\infty}^{\infty} \mathrm{d}x_1 \dots \mathrm{d}x_n \exp\left(\mathrm{i} \sum_j (x_j - \mu_j \alpha) z_j - \frac{1}{2} (X_\mu)^T R^{-1} (X_\mu)\right). \tag{3}$$

The integration contour C is along the real axis, except near the origin where it crosses the imaginary axis in the lower half-plane. Using a well-known integration formula (see, e.g., Ref. 8), Eq. (3) gives

$$Q = \frac{1}{(2\pi \mathrm{i})^n} \sum_{\mu_i=\pm 1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{\mathrm{d}z_1 \ldots \mathrm{d}z_n}{z_1 \ldots z_n} \times$$

$$\exp\left(-\mathrm{i}\alpha \sum_j \mu_j z_j - \frac{1}{2} \sum_{ij} \mu_i \mu_j z_i z_j r_{ij}\right). \tag{4}$$

Integrals over z variables can be performed if the exponential function in the integrand is expanded in powers of the correlation matrix elements r_{ij} and the level α. In the lowest order we obtain

$$Q = \frac{1}{2^n} \sum_\mu \left(1 - \alpha\sqrt{\frac{2}{\pi}} \sum_j \mu_j + \frac{2}{\pi} \sum_{i<j} r_{ij}\mu_i\mu_j + \ldots\right) \approx$$

$$\frac{1}{2^n} \sum_\mu \exp\left(-\alpha\sqrt{\frac{2}{\pi}} \sum_j \mu_j + \frac{2}{\pi} \sum_{i<j} r_{ij}\mu_i\mu_j\right). \tag{5}$$

Comparing this results to the Ising model, we see that in the lowest order the ensemble of surfaces obtained by a level cut of Gaussian random field is equivalent to that of the Ising model with the interaction $\beta V(r_{ij}) = -2r_{ij}/\pi$. The range of the interaction is thus determined by the range of the covariance function of the generating Gaussian field r_{ij}. The first order approximation is accurate if r_{ij} is small for all i and j.

Symmetrical states obtained by the level cut at $\alpha = 0$ correspond to the case of no external magnetic field. For deviation from symmetry, the model corresponds to the Ising model in the external magnetic field given by $\beta H = \alpha\sqrt{2/\pi}$.

We are now in a position to understand how the accuracy in specifying the position of level cuts affects the calculation of the system entropy. The behaviour of the system depends on the ratio of the two length scales: of the range of correlations of the Gaussian field (i.e. the range of the "potential" $V(r_{ij})$) to the lattice constant. When r_{ij} is small for nearest neighbours i and j, the system behaves like an Ising model with short-range interaction. The entropy is much smaller than that of the generating field.

On the other hand, if the range of the covariance function of the generating field is large compared to the lattice constant (i.e., the accuracy in the definition of the surface, or the sampling rate of continuous signals), we have the Ising model with long–range interaction. From the work of Baker [8], we know that such Ising models approach the Gaussian mean field behaviour with the deviations being of the order $1/R$, where R is the number of spins in the range of the interaction. The entropy of the levelled Gaussian field is then similar to that of the generating field.

The model also indicates that non-symmetrical microemulsion or L_3 phase states have a lower entropy than the symmetrical states. In describing such states, one must not use the usual simple expression for the entropy of a random Gaussian field [9], but rather use more complicated expressions corresponding the decreased entropy obtained when an Ising system is placed into an external magnetic field.

The previous observation relates to some well-known empirical findings in the field of signal representation. When images are reconstructed from data obtained by a single level cut, much better results are obtained if the level cut is selected so that the resulting black/white image is approximately symmetrical [5]. Although the number of equations used to reconstruct an image in non–symmetrical cases can be arbitrarily large and exact reconstruction should be possible, the process produces poor results. This is due the finite precision of the data (corresponding to the Ising model lattice parameter) and a lower information content, corresponding to the entropy of the Ising model in a magnetic field.

At present, the argument is approximate: First order results become progressively worse as the interaction range or the asymmetry increase (the errors become large as $r_{ij} \to 1$), and more accurate calculation appears difficult because in the higher–order terms, many lattice sites become coupled.

4. *Link to liquid state physics*

The problem of calculating the entropy of disordered surfaces, or that of sampled and digitised signals, may be approached using the formalism developed in calculations of the entropy of liquids with long–range interaction between molecules. Microemulsions and L_3 phases may be considered as two-component fluids, consisting of "oil" and "water" or "inside" and "outside" sites. When signal is digitised into n levels, we have a correspondence to an n-component fluid. The analogy is exact for L_3 phases (where there is water on both "inside" and "outside") or for symmetrical signals. In the case of microemulsions, we assume that the short–range part of the entropy, corresponding to the precise molecular arrangement over several diameters, is not affected by the oil/water partitioning.

The entropy of a multicomponent fluid may be expressed as a functional of its correlation functions. It can, therefore, be calculated for a variety of stochastic ensembles of functions, as long as the correlation functions are accessible. The method is not restricted to Gaussian statistics, and can be applied to a variety of signal statistics. It will be easier to find good approximations when the correlation range is short and when the first nonvanishing cumulants are small.

In the case of levelled Gaussian field, correlation functions can be evaluated in terms of the covariance function of the generating field. Particularly simple results are obtained for the symmetrical case. The pair correlation functions are evaluated by integration over the bivariate Gaussian distribution [10] as, e.g. $g_{++}(ij) = 1 + \frac{2}{\pi} \arcsin r_{ij}$. Third order correlation functions, which will be needed in entropy calculation are just a simple generalisation. One finds, e.g.

$$g_{+++}(ijk) = 1 + \frac{2}{\pi}(\arcsin r_{ij} + \arcsin r_{ik} + \arcsin r_{jk}). \tag{6}$$

Fourth–order correlation functions are no longer simple. More general correlation functions for unsymmetrical cases have been studied by Roberts and Teubner [11].

The effective interaction is long-ranged, and we must use appropriate expressions for the entropy. A resummation of ring diagrams, as used in the hypernetted chain approximation, has been performed by Hernando [12]. In a recent study, Laird, Wang and Haymet [13] generalised that result to mixture fluids and demonstrated its accuracy in the case of Coulomb interaction. Although the resulting expression is rather complicated, the only inputs into the calculation are second and third order correlation functions of the mixed fluid.

The accuracy of the standard liquid state approximations for the entropy of an ensemble of surfaces, obtained by level cutting, remains to be evaluated. The random Gaussian field before the levelling operation, where exact entropy is known, corresponds to a fluid with an infinite number of components. Application of approximate expression to this case should provide information on the accuracy of various approximation that use only low–order correlation functions for evaluation of the entropy.

5. *Discussion*

A wide variety of structures, typical of phase-separated materials or surfaces, can be described with the morphologies obtained from the levelling operation on random Gaussian fields. The variability comes from the choice of the spectral density of the generating field, where suitable periodicity as well as the characteristic range of the interaction can be accommodated. Understanding of the phase and dilution behaviour of equilibrium phases described by such structures requires the knowledge of the corresponding entropy.

We considered here two different approaches to the problem of evaluating the entropy, or the information content of the states, obtained by levelling of the random generating field. The analogy with the Ising model indicates that the information loss is reciprocally proportional to the number of lattice sites within the range of the correlations in the generating field.

In the case of microemulsions or L_3 phases, this means that the entropy of the generating field provides a good approximation to the entropy of surfaces, as long as the correlation length of the field is large in comparison with the accuracy in the definition of the surfaces. Entropy of more dilute structures will be better approximated by the entropy of the generating field. Of course, the random Gaussian field topology is only appropriate within a certain well-defined range of values of membrane elastic constants, as shown in a recent simulation [14].

The question of the information content of a function obtained by zero-crossing, or more generally level crossing, arises in many areas of signal processing. Proposed developments should provide an accurate measure of information loss in commonly used representations of continuous signals. Similar considerations arise in the field of vision research. The present work indicates how the information content of sampled and digitised signals depends on the correlation range of the original signal.

Acknowledgement

I am grateful to P. Pieruschka for many discussions during the initial stage of this work and to S. Safran for a preprint of Ref. 9.

References

1) N. F. Berk, Phys. Rev. Lett. **58** (1987) 2718;
2) M. Teubner, Europhys. Lett. **14** (1991) 403;
3) P. Pieruschka and S. Marčelja, J. Phys. II (France) **2** (1992) 235;
4) P. Pieruschka and S. Safran, Europhys. Lett. **22** (1993) 625;
5) A. Zakhor and A. Oppenheim, Proc. IEEE **78** (1990) 31;
6) D. E. Knuth, ACM Trans. Graphics **6** (1987) 245;
7) C. E. Shannon, in C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana. Ill. 1949;
8) G. A. Baker, Phys. Rev. **126** (1962) 2071;
9) P. Pieruschka and S. Safran, Europhys. Lett. **31** (1995) 207;
10) W. F. Sheppard, Phil. Trans. **A 192** (1898) 101; Proc. Roy. Soc. **62** (1898) 170;
11) A. P. Roberts and M. Teubner, Phys. Rev. E **51** (1995) 4141;
12) J. A. Hernando, Mol. Phys. **69** (1990) 327;
13) B. B. Laird, J. Wang and A. D. J. Haymet, Phys. Rev. E **47** (1993) 2491; B. B. Laird and A. D. J. Haymet, J. Chem. Phys. **100** (1994) 3775;
14) P. Pieruschka and S. Marčelja, Langmuir **10** (1994) 345;

MIKROEMULZIJE, L_3 FAZE I GUBITAK INFORMACIJA U SIGNALNOJ REPREZENTACIJI

Plohe dvojno–neprekidnih neuređenih mikroemulzijskih struktura ili L_3 faza mogu se opisati kao razinski rezovi Gaussovog polja izvjesnog spektra. Razmatramo razliku između entropije tih stanja i dobro poznate entropije generirajućeg nasumnog polja. U drugom smislu, ta je razlika jednaka smanjenju Shannonovog sadržaja informacija koji slijedi iz procesa odabira i digitalizacije neprekidnih stohastičkih signala.