# Pairwise Markov Random Fields and its Application in Textured Images Segmentation

Wojciech Pieczynski and Abdel-Nasser Tebbache

*Département Signal et Image*
*Institut National des Télécommunications,*
*9, rue Charles Fourier, 91000 Evry, France*

*E-mail Wojciech.Pieczynski@int-evry.fr*

## Abstract

*The use of random fields, which allows one to take into account the spatial interaction among random variables in complex systems, is a frequent tool in numerous problems of statistical image processing, like segmentation or edge detection.*

*In statistical image segmentation, the model is generally defined by the probability distribution of the class field, which is assumed to be a Markov field, and the probability distributions of the observations field conditional to the class field. In such models the segmentation of textured images is difficult to perform and one has to resort to some model approximations. The originality of our contribution is to consider the markovianity of the pair (class field, observations field). We obtain a different model; in particular, the class field is not necessarily a Markov field. The model proposed makes possible the use of Bayesian methods like MPM or MAP to segment textured images with no model approximations. In addition, the textured images can be corrupted with correlated noise. Some first simulations to validate the model proposed are also presented.*

## 1. Introduction

We propose in this paper a new Pairwise Markov Random Field model (PMRF), which is original with respect to the classical Hidden Markov Random Field model (HMRF) [ChJ93]. The main difference is that in a PMRF the class field is not necessarily a Markov field. One advantage, which will be developed in the following, is that textured images can be segmented without any model approximation.

To be more precise, let $S$ be the set of pixels, $X = (X_s)_{s \in S}$ the random field of classes, and $Y = (Y_s)_{s \in S}$ the random field of observations. The problem of segmentation is then the problem of estimating the realizations of the field $X$ from the observations of the realizations of the field $Y$. Prior to considering any statistical segmentation method, one must define the probability distribution $P_{(X,Y)}$ of the random field $(X,Y)$.

In the classical HMRF model this distribution is given by the distribution of $X$, which is a Markov field, and the set $P_Y^{X=x}$ of the distributions of $Y$ conditional to $X = x$. Under some assumptions on $P_Y^{X=x}$, the posterior distribution $P_X^{Y=y}$ of $X$ is a Markov distribution and different Bayesian segmentation techniques like MPM [MMP87], MAP [GeG84] , or ICM [Bes86] can be applied. These assumptions can turn out to be difficult to justify when dealing with textured images, as detailed in [DeE87], [KDH88], [WoD92].

The idea of the PMRF model we propose is to consider directly the Markovianity of the pairwise random field $Z = (X,Y)$. The distribution of $X$ is then the marginal distribution of $P_{(X,Y)}$ and thus it is not necessarily a Markovian distribution. What counts is that $P_X^{Y=y}$ remains Markovian and so different Bayesian segmentations specified above can be used. When the particular problem of segmenting textured images with correlated noise is considered, the PMRF model is quite suitable because different Bayesian segmentation methods above can be applied without any model approximation.

The organization of the paper is following. In the next section we specify the difference between HMRF and PMRF in the case of simple Gaussian noise fields. Section 3 is devoted to some simulations and some concluding remarks are presented in section 4.

## 2. Pairwise Markov Random Field and textured image segmentation

### 2.1 Simple case of Hidden Markov Field

Let us consider the set of pixels $S$ and $X = (X_s)_{s \in S}$ a random field, each random variable $X_s$ taking its values in the class set $\Omega = \{\omega_1, \omega_2\}$. The field $X$ is Markovian with respect to four nearest neighbors if its distribution is written

$$P[X=x]=\lambda\exp\left[-\sum_{(s,t)\text{neighbors}}\varphi_1(x_s,x_t)-\sum_s\varphi_2(x_s)\right] \quad (2.1)$$

where "$(s,t)$ neighbors" means that the pixels $s$ and $t$ are neighbors and lie either on a common row or on a common column. The random field $Y=(Y_s)_{s\in S}$ is the field of observations and we assume that each $Y_s$ takes its values in $R$. The distribution of $(X,Y)$ is then defined by (2.1) and the distributions of $Y$ conditional on $X=x$. Assuming that the random variables $(Y_s)$ are independent conditionall to $X$ and that the distribution of each $Y_s$ conditional to $X=x$ is equal to its distribution conditional to $X_s=x_s$, we have :

$$P[Y=y|X=x]=\prod_s f_{x_s}(y_s) \quad (2.2)$$

where $f_{x_s}$ is the density of the distribution of $Y_s$ conditional to $X_s=x_s$. Thus :

$$P[X=x,Y=y]= \\ =\lambda\exp[-\sum_{(s,t)\text{neighbors}}\varphi_1(x_s,x_t)-\sum_s[\varphi_2(x_s)+Log f_{x_s}(y_s)]] \quad (2.3)$$

So the pairwise field $(X,Y)$ distribution is Markovian and the distribution of $X$ conditional to $Y=y$ is still Markovian. It is then possible to simulate realizations of $X$ according to its distribution conditional to $Y=y$, which affords the use of Bayesian segmentation techniques like MPM or MAP.
In practice, the random variables $(Y_s)$ are not, in general, independent conditionally on $X$. In particular, (2.2) is too simple to allow one to take texture into account. For instance, if we consider that texture is a Gaussian Markov random field realization [CrJ83], (2.2) should be replaced with :

$$P[Y=y|X=x]= \\ =\lambda(x)\exp[-\sum_{(s,t)\text{neighbors}}a_{x_sx_t}y_sy_t-\frac{1}{2}\sum_s[a_{x_sx_s}y_s^2+b_{x_s}y_s]] \quad (2.4)$$

The field $Y$ is then Markovian conditionally on $X$, which models textures. The drawback is that the product of (2.1) with (2.4) is not, in general, a Markov distribution. In fact, for the covariance matrix $\Gamma(x)$ of the Gaussian distribution of $Y=(Y_s)_{s\in S}$ (conditional to $X=x$), we have :

$$\lambda(x)=[(2\pi)^N\det(\Gamma(x))]^{-1/2} \quad (2.5)$$

which is not, in general, a Markov distribution with respect to $x$.
Finally, $X$ is Markovian, $Y$ is Markovian conditionally on $X$, but neither $(X,Y)$, nor $X$ conditionally on $Y$, are Markovian in general. This lack of the posterior Markovianity invalidates the rigorous application of MPM or MAP.

## 2.2 Simple case of Pairwise Markov Field

To circumvent the difficulties above we propose to consider the Markovianity of $(X,Y)$. Specifically, we put

$$P[X=x,Y=y]= \\ =\lambda\exp[-\sum_{(s,t)\text{neighbors}}\varphi[(x_s,y_s),(x_t,y_t)]-\sum_s\varphi^*[(x_s,y_s)]]= \\ =\lambda\exp[-\sum_{(s,t)\text{neighbors}}[\varphi_1(x_s,x_t)+a_{x_sx_t}y_sy_t+b_{x_sx_t}y_s+ \\ +c_{x_sx_t}y_t]-\sum_s[\varphi_2(x_s)+a_{x_sx_s}y_s^2+b_{x_s}y_s]] \quad (2.6)$$

The Markovianity of the pairwise field $(X,Y)$ implies the Markovianity of $Y$ conditionally on $X$, and the Markovianity of $X$ conditionally on $Y$. The first property allows one to model textures, as in (2.4), and the second one makes possible to simulate $X$ according to its posterior distribution, which allows us to use Bayesian segmentation methods like MPM or MAP.
Let us briefly specify how to simulate realizations of the pair $(X,Y)$. The pair $(X,Y)$ being Markovian, we can specify the distribution of each $(X_s,Y_s)$ conditionally on its neighbors. Let us consider the calculus of the distribution of $(X_s,Y_s)$ conditional to the four nearest neighbors :

$$[(X_{t_1},Y_{t_1}),(X_{t_2},Y_{t_2}),(X_{t_3},Y_{t_3}),(X_{t_4},Y_{t_4})]= \\ [(x_{t_1},y_{t_1}),(x_{t_2},y_{t_2}),(x_{t_3},y_{t_3}),(x_{t_4},y_{t_4})] \quad (2.7)$$

This distribution can be written as

$$h(x_s,y_s)=p(x_s)f_{x_s}(y_s) \quad (2.8)$$

where $p$ is a probability on the set of classes and, for each class $x_s$, $f_{x_s}$ is the density of the distribution of $Y_s$ conditional to $X_s=x_s$ ($p$ and $f_{x_s}$ also depend on $(x_{t_1},y_{t_1}),(x_{t_2},y_{t_2}),(x_{t_3},y_{t_3}),(x_{t_4},y_{t_4})$, which are fixed in the following and so will be omitted). (2.8) makes the sampling of $(X_s,Y_s)$ quite easy: one samples $x_s$ according to $p$, and then $y_s$ according to $f_{x_s}$. We thus have:

$$P\{(X_s,Y_s)=(x_s,y_s)| \\ |[(X_{t_1},Y_{t_1}),\ldots,(X_{t_4},Y_{t_4})]=[(x_{t_1},y_{t_1}),\ldots,(x_{t_4},y_{t_4})]\} \\ \propto\exp[-\sum_{i=1,\ldots,4}\varphi[(x_s,y_s),(x_{t_i},y_{t_i})]-\varphi^*[(x_s,y_s)]]= \\ =\exp[-\sum_{i=1,\ldots,4}[\varphi_1(x_s,x_{t_i})+a_{x_sx_{t_i}}y_sy_{t_i}+b_{x_sx_{t_i}}y_s+ \\ +c_{x_sx_{t_i}}y_{t_i}]-[\varphi_2(x_s)+a_{x_sx_s}y_s^2+b_{x_s}y_s]] \quad (2.9)$$

This can be written $h(x_s, y_s) = p(x_s) f_{x_s}(y_s)$, where $f_{x_s}$ is a Gaussian density defined by the following mean $M_{x_s}$ and variance $\sigma^2_{x_s}$ :

$$M_{x_s} = -\frac{b_{x_s} + \sum_{i=1,\dots,4} (a_{x_s x_{t_i}} y_{t_i} + b_{x_s x_{t_i}})}{2a_{x_s x_s}}, \quad \sigma^2_{x_s} = \frac{1}{2a_{x_s x_s}} \quad (2.10)$$

and $p$ the probability given on the set of classes with :

$$p(x_s) \propto \sqrt{(a_{x_s x_s})^{-1}} \exp[\frac{(b_{x_s} + \sum_{i=1,\dots,4} a_{x_s x_{t_i}} y_{t_i} + b_{x_s x_{t_i}})^2}{4a_{x_s x_s}} - \quad (2.11)$$
$$-\varphi_2(x_s) - \sum_{i=1,\dots,4} (\varphi_1(x_s, x_{t_i}) + c_{x_s x_{t_i}} y_{t_i})]$$

Finally, the main differences between the classical HMRF model and the new PMRF we propose are :

(i) The distribution of $X$ (its prior distribution) is Markovian in HMRF and is not necessarily Markovian in PMRF;

(ii) The posterior distribution of $X$ is not necessarily Markovian in HMRF and is Markovian in PMRF;

(iii) In the case of case of images which are textured, and possibly corrupted with correlated noise, the PMRF allows one to apply the Bayesian MPM or MAP methods without any model approximations.

**Remarks**

1. The classical HMRF can also be applied in the case of multisensor images [YaG95]. For $m$ sensors the observations on each pixel $s \in S$ are then assumed to be a realization of a random vector $Y_s = \left[Y_s^1, \dots, Y_s^m\right]$. It is possible to consider a multisensor PRMF. For example, a multisensor PMRF would be obtained by replacing in (2.6) $y_s y_t$ and $y_s^2$ by $\phi_1\left[(y_s^1, \dots, y_s^m), (y_t^1, \dots, y_t^m)\right]$ and $\phi_2\left[(y_s^1, \dots, y_s^m)\right]$.

2. In some particular cases the classical model allows one to take correlated noise into account [Guy93], [Lee98]. The difficulties arise when wishing to consider the cases when correlations vary with classes.

### 2.3 General case of PMRF

Generalizing of the distribution given by (2.6) does not pose any problem. Let us consider $k$ classes $\Omega = \{\omega_1, \dots, \omega_k\}$, $m$ sensors (each $Y_s = (Y_s^1, \dots, Y_s^m)$ takes its values in $R^m$), and $C$ a set of cliques defined by some neighborhood system. The random field $Z = (Z_s)_{s \in S}$, with $Z_s = (X_s, Y_s)$, is a Pairwise Markov Random Field if its distribution may be written as

$$P[Z = z] = \lambda \exp[-\sum_{c \in C} \varphi^c(z_c)] \quad (2.12)$$

Let us note that the existence of the distribution (2.12) is not ensured in a general case.
In particular, three sensor PMRF can be used to segment colour images.

**Remark**

As mentioned above, $X$ is not necessarily Markovian in a PMRF. This could be felt as a drawback, because the distribution of $X$ models the "prior" , i.e., without any observation, knowledge we have about the class image. Of course, this "prior" distribution of $X$ also exists in PMRF (it is the marginal distribution of $P_{(X,Y)}$), but it is not Markovian. The gravity of this "drawback" is undoubtedly difficult to discuss in the general case. However, supposing that it really is a drawback, let us mention that it also exists in the classical HMRF model. In fact, even in the very simple case defined by (2.3), the observation field $Y$ is not a Markov field. So, one could consider that the non Markovianity of $X$ in the PMRF model is not stranger that the non Markovianity of $Y$ in the classical HMRF model.

## 3. Visual examples

We present in this section two simulations and two segmentations by the MPM. We consider rather noisy cases : one can hardly distinguish the class image in the noisy one. One can notice that although the class images are not Markov fields realizations, they look like.
The two realizations of PMRF presented in Fig.1 are Markovian with respect to four nearest neighbors; the distribution of $(X,Y)$ is written :

$$P[X = x, Y = y] =$$
$$= \lambda \exp[-\sum_{(s,t)\text{neighbors}} \varphi[(x_s, y_s), (x_t, y_t)] - \sum_s \varphi^*[(x_s, y_s)]] \quad (3.1)$$

with

$$\varphi[(x_s, y_s), (x_t, y_t)] = \frac{1}{2}(a_{x_s x_t} y_s y_t + b_{x_s x_t} y_s + c_{x_s x_t} y_t + d_{x_s x_t}) \quad (3.2)$$
$$\varphi^*[(x_s, y_s)] = \frac{1}{2}(\alpha_{x_s} y_s^2 + \beta_{x_s} y_s + \gamma_{x_s x_t})$$

The different coefficients in (3.2) are given in Tab.1. Let us notice that it is interesting, in order to have an idea about the noise level, to dispose of some information about the distributions of $Y$ conditional to $X = x$. In fact, these are

Gaussian distributions and knowing some parameters like means and variances can provide some information about the noise level. Of course, the noise level also depends on different correlations and the prior distribution of $X$.

We may note that some of the coefficients in (3.2) are simply linked with means or variances of the distributions of $Y$ conditional to $X = x$. In fact, denoting by $\Sigma_x$ the covariance matrix of the Gaussian distribution of $Y$ conditional to $X = x$ and putting $Q_x = [q_{st}^x]_{s,t\in S} = \Sigma_x^{-1}$, we have :

$$P[Y = y|X = x] \propto \exp\left[-\frac{(y - m_x)^t Q_x (y - m_x)}{2}\right] \qquad (3.3)$$

Developing (3.3) and identifying to (3.2) we obtain

$$m_{x_s} = -\frac{\beta_{x_s}}{2\alpha_{x_s}}, \qquad \sigma_{x_s}^2 = \frac{1}{\alpha_{x_s}} \qquad (3.4)$$

So, all other parameters being fixed, one can use (3.4) to vary the noise level. For instance, keeping the same variances the noise level increases when one makes the means approach each other. Otherwise, there are no simple links between correlations of the random variables $(Y_s)$ (conditionally on $X = x$) and the coefficients in (3.2). The correlations in Table 1, whose variations make appear different textures, are estimated ones. The values of the means show that the level of the noise is rather strong, which is confirmed visually.

Image 1

Image 4

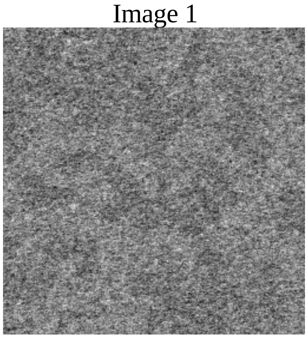

Image 2

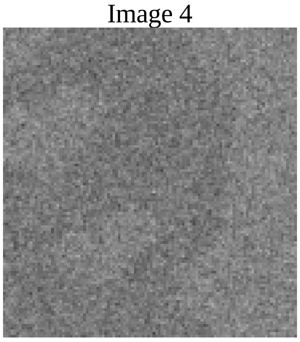

Image 5

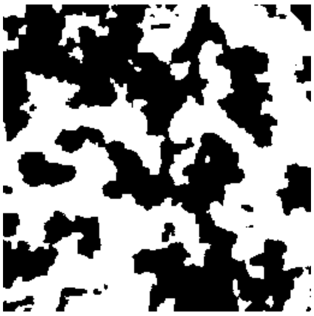

Image 3

Image 6

Fig. 1. Two realizations of Pairwise Markov Fields (Image 1, Image 2), (Image 4, Image 5), and the MPM segmentations of Image 2 (giving Image 3), and Image 5 (giving Image 6), respectively.

| | Images 1, 2, 3 | Images 4, 5, 6 |
|---|---|---|
| $\alpha_{x_s}$ | 1 | 1 |
| $\beta_{x_s}$ | $-2m_{x_s}$ | $-2m_{x_s}$ |
| $\gamma_{x_s x_t}$ | $m_{x_s}^2$ | $m_{x_s}^2$ |
| $a_{x_s x_t}$ | $-0,4$ | $-0,1$ |
| $b_{x_s x_t}$ | $-0,4m_{x_t}$ | $-0,1m_{x_t}$ |
| $c_{x_s x_t}$ | $-0,4m_{x_s}$ | $-0,1m_{x_s}$ |
| $d_{x_s x_t}$ | $0,4m_{x_s}m_{x_t} + \varphi(x_s, x_t)$ | $-0,1m_{x_s}m_{x_t} + \varphi(x_s, x_t)$ |
| $m_1$ | $-0,3$ | 1 |
| $m_2$ | $0,3$ | $1,5$ |
| $\sigma_1^2$ | 1 | 1 |
| $\sigma_2^2$ | 1 | 1 |
| $\rho_{11}$ | $0,26$ | $0,05$ |
| $\rho_{22}$ | $0,26$ | $0,07$ |
| $\tau$ | 13,1% | 07,9% |
| $Nb$ | $30\times30$ | $30\times30$ |

Tab.1

$\alpha_{x_s}$, ... , $d_{x_s x_t}$ : functions in (3.2), the function $\varphi$ being defined by $\varphi(x_s, x_t) = -1$ if $x_s = x_t$ and $\varphi(x_s, x_t) = 1$ if $x_s \neq x_t$. $m_1$, $m_2$ , $\sigma_1^2$, $\sigma_2^2$ : the means and the variances in (3.3). $\rho_{11}$, $\rho_{22}$ : the estimated covariances inter-class (neighboring pixels). $\tau$ : the error rate of wrongly classified pixels with MPM. $Nb = n_1 n_2$ : the number of iterations in MPM (the posterior marginals estimated from $n_1$ realizations, each realization obtained after $n_2$ iterations of the Gibbs Sampler).

## 4 Conclusions

We proposed in this paper an novel model called Pairwise Markov Random Field (PMRF). A random field of classes $X$ and a random field of observations $Y$ form a PMRF when the pairwise random field $Z = (X, Y)$ is a Markov field. Such a model is different from the classical Hidden Markov Random Field (HMRF); in particular, in PMRF the random field $X$ is not necessarily a Markov field.

The PMRF allows one to deal with the statistical segmentation of textured images which can be, in addition, corrupted with correlated noise. On the contrary to the use of hierarchical models [DeE87], this can be done in the framework of the model, without any approximations. Roughly speaking, in the Hierarchical HMRF the prior distribution of $X$ is Markovian and its posterior distribution is not Markovian; and in PMRF the prior distribution of $X$ is not Markovian and its posterior distribution is Markovian. When using a Bayesian method of segmentation like MPM or MAP we have to make some approximations when using Hierarchical HMRF, and we have not when using PMRF. Furthermore, the distributions of $Y$ conditional to $X$, which model different textures and different possibly correlated noises, can be strictly the same in the both Hierarchical HMRF and PMRF models.

We have presented two simulations of PMRF and two results of the Bayesian MPM segmentation of the observation fields. The two cases presented are rather noisy and the results show that the well known efficiency of the HMRF can also occur when using PMRF.

As perspectives for further work, let us mention two important points. First, the existence of the distribution given by (2.12) is not ensured in the general case. Even in the simple Gaussian case given by (2.6) we should verify that all Gaussian distributions of $Y$ conditionally on $X$ exist. There exist some conditions of existence of Gaussian fields [Guy95] and thus one possible way of verifying the existence of PMRF could be the verification of the existing conditions "uniformly" with respect to $X$. The second problem is the parameter estimation one. One could view applying the general Iterative Condition Estimation (ICE [Pie92]), which gives satisfying results in some classical situations [DMP97], [GiP97], [SaP97]. Using ICE requests considering an estimator from complete data $(X, Y)$: one possible way of seeking such an estimator could be considering the stochastic gradient [You88], applied to $(X, Y)$ instead of $X$.

*Acknowledgment. We thank Alain Hillion, Directeur Scientifique de l'École Nationale Supérieure des Télécommunications de Bretagne, for numerous discussions which greatly helped the writing of this paper.*

## 5. References

[Bes86] J. Besag, On the statistical analysis of dirty pictures, *Journal of the Royal Statistical Society*, Series B, 48, pp. 259-302, 1986.

[ChJ93] R. Chellapa, A. Jain Ed., Markov Random Fields, Theory and Application, *Academic Press, San Diego,* 1993.

[CrJ83] G. R. Cross and A. K. Jain, Markov Random Field Texture Models, *IEEE Trans. on PAMI*, Vol. 5, No. 1, pp. 25-39, 1983.

[DMP97] Y. Delignon, A. Marzouki, and W. Pieczynski, Estimation of Generalized Mixture and Its Application in Image Segmentation, *IEEE Transactions on Image Processing,* Vol. 6, No. 10, pp. 1364-1375, 1997.

[DeE87] H. Derin and H. Elliot, Modelling and segmentation of noisy and textured images using Gibbs random fields, *IEEE Trans. on PAMI*, Vol. 9, No. 1, pp. 39-55, 1987.

[GeG84] S. Geman, G. Geman, Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, *IEEE Trans. on PAMI*, Vol. 6, No. 6, pp. 721-741, 1984.

[GiP97] N. Giordana and W. Pieczynski, Estimation of Generalized Multisensor Hidden Markov Chains and Unsupervised Image Segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 19, No. 5, pp. 465-475, 1997.

[Guy95] X. Guyon, Random Fields on Network. Modeling, Statisticsz, and Applications. Springer-Verlag, Probabily and its Applications, New York, 1995.

[KDH88] P. A. Kelly, H. Derin, and K.D. Hart, Adaptive segmentation of Speckled images using a hierarchical random field model, *IEEE Trans. on ASSP,* Vol. 36, No 10, pp. 1628-1641, 1988.

[LaD89] S. Lakshmanan, H. Derin, Simultaneous parameter estimation and segmentation of Gibbs random fields, *IEEE Transactions on PAMI*, Vol. 11, pp. 799-813, 1989.

[Lee98] T. C. M. Lee, Segmenting images corrupted by correlated noise, *IEEE Trans. on PAMI*, Vol. 20, No. 5, pp. 481-492, 1998.

[MMP87] J. Marroquin, S. Mitter, T. Poggio, Probabilistic solution of ill-posed problems in computational vision, *Journal of the American Statistical Association*, 82, pp. 76-89, 1987.

[Pie92] W. Pieczynski, Statistical image segmentation, *Machine Graphics and Vision*, Vol. 1, No. 1/2, pp. 261-268, 1992.

[SaP97] F. Salzenstein and W. Pieczynski, Parameter Estimation in Hidden Fuzzy Markov Random Fields and Image Segmentation, *Graphical Models and Image Processing,* Vol. 59, No. 4, pp. 205-220, 1997.

[WoD92] C. S. Won and H. Derin, Unsupervised segmentation of noisy and textured images using Markov random fields, *CVGIP: Graphical Models and Image Processing*, Vol. 54, No 4, 1992, pp. 308-328.

[YaG95] T. Yamazaki and D. Gingras, Image classification using spectral and spatial information based on MRF models, *IEEE Trans. on IP,* Vol. 4, No. 9, pp. 1333-1339, 1995.

[You88] L. Younes, Estimation and annealing for Gibbsian fields, *Annales de l'Institut Henri Poincaré*, Vol. 24, No. 2, 1988, pp. 269-294.