

MODEL MANIFOLDS FOR ANALYSING DATA OF WAVE SPLITTING

E. VAN GROESEN[†] AND W.L. IJZERMAN

University of Twente, Department of Applied Mathematics
P.O. Box 217, 7500 AE Enschede, The Netherlands

Abstract. When large amounts of data are available, such as from numerical calculations of a physical-technical process, insight in the underlying process can be gained if the data are used to specify adjustable parameters in a preliminary model of the phenomenon. Information from the signal can thus be obtained that would otherwise be difficult to detect. A general approach to the modelling of data is presented in this paper, and is illustrated for numerical data of the splitting of a wave due to bottom variations.

1. Introduction and motivation

In analysing a spatial-temporal signal, the *method of principal components*, also called POD: proper orthogonal decomposition, is often used nowadays (Aubry *et al*, 1991; Berkooz *et al*, 1993; Sirovich *et al*, 1989). This method has proven to be useful in various cases when the signal is generated by a system with a (low-dimensional) attractor; the spatial components determine the geometry of the attractor, while the temporal components describe the evolution on the attractor.

This method uses the signal as its only input: no knowledge of the underlying system is required. If some (qualitative) knowledge of the system is known, it can be expected that this knowledge can be used profitable in the analysis of the signal. In particular the signal can be used in the "calibration" of parameters that appear in an adjustable preliminary model:

[†]Part of the research is sponsored by the Commission of the European Communities, Directorate General XII-B, Joint Research Project CII*-CT93-0018, and by the Royal Netherlands Academy of Arts and Sciences in the programma for scientific cooperation Netherlands-Indonesia.

the signal will determine the evolution of the parameters.

To be more specific, we start to review briefly the method of principal components.

Let the signal be given by $s(x, t)$ where x is the spatial variable and t the time; it is no restriction to assume that s is a scalar quantity. Then N spatial components (topos) $\phi_k(x), k = 1 \dots N$ and temporal components (chronos) $a_k(t), k = 1 \dots N$ are sought from the minimization problem

$$\text{Min } \left\{ \int \int [s(x, t) - \sum a_k(t) \phi_k(x)]^2 dx dt \mid a_k(t), \phi_k(x), k = 1 \dots N \right\}.$$

The spatial components are normalized, and additional conditions like boundary conditions may be imposed.

The solution of this minimization problem, say $\sum a_k^*(t) \phi_k^*(x)$, is then the (N -dimensional) representation of the signal.

Remarks

1. The components can also be found in another way: the spatial (temporal) components are the eigenfunctions of the integral operator with the temporal (spatial) correlation function as kernel. For instance, the first spatial component ϕ_1 is the solution of the maximization problem

$$\text{Max } \left\{ \int \int C(x, y) \phi(x) \phi(y) dx dy \mid \int \phi^2(x) dx = 1 \right\} \quad (1.1)$$

with C the temporal correlation function

$$C(x, y) = \int s(x, t) s(y, t) dt \quad (1.2)$$

The maximizing characterization gives the interpretation of $\phi_1^*(x)$ as being the time-averaged main spatial structure that is present in the signal. The corresponding temporal component $a_1^*(t)$ is the projection of the signal along ϕ_1^* : $a_1^*(t) = \int s(x, t) \phi_1^*(x) dx$.

2. The other spatial components can be found in a successive way by maximizing over the orthogonal complement. These components characterize in decreasing order of importance the remaining spatial structures in the signal.
3. The number N that is required to represent the signal in a sufficient degree of accuracy (for instance to capture 95% of the energy) is an indication of the complexity of the signal and, under some ergodicity properties, of the complexity of the attractor of the underlying system; the spatial components then represent the attractor.

For the following it is useful to interpret the method of principal components as the search for spatial components ϕ_k that are such that the signal $s(x, t)$ can be represented as good as possible by a trajectory $t \mapsto u(t)$ in their linear span

$$M = \left\{ u = \sum_{k=1}^N a_k \phi_k(x) \mid a_k \in \mathcal{R} \right\}.$$

The measure to determine the optimal trajectory is the spatial-temporal L_2 -norm (denoted by $||| \dots |||$) of the difference of the signal from the trajectory:

$$\text{Min } \{ |||s - u(t)||| \mid u(t) \in M \}.$$

This interpretation allows for a simple modification when some information can be extracted a priori from the underlying system, or, more generally, when some qualitative information is known about the signal. In fact, suppose that a manifold \mathcal{M} can be defined consisting of (spatial) structures that are expected to be present in the signal dominantly; then a trajectory in \mathcal{M} will be sought to represent the signal by minimizing some distance function between the trajectory and the signal.

Both the choice of the manifold \mathcal{M} and the choice of distance function are essential in the resulting representation of the signal. The manifold \mathcal{M} will be called the *model manifold*; the choice of the distance function is equivalent to specifying a projection onto this manifold.

In the next section we present a simple example to motivate the necessity of looking for modifications of the principal component method described above. The example is a single wave travelling at constant speed; we propose a modification that determines both the profile and the speed. In section 3 we consider the much more difficult case of a signal that describes the splitting of one wave into two waves with different amplitudes and speeds. In the final section we make some remarks and formulate conclusions.

2. Reconstruction of a travelling wave

Consider a signal of a travelling wave of permanent wave profile f that is translated with a fixed speed c :

$$s(x, t) = f(x - ct). \quad (2.3)$$

Analytic consideration

First assume that the signal is known for all x and t ; then the temporal correlation function becomes a function of the difference $C(x, y) = C_0(x - y)$, and the corresponding integral operator is a convolution: the eigenvalue

problem for the spatial principal components becomes

$$\int C_0(x-y)\phi(y)dy = \mu\phi(x),$$

or after Fourier transformation (denoted by hats)

$$\hat{C}_0(k)\hat{\phi}(k) = \mu\hat{\phi}(k).$$

This leads to a continuous spectrum with harmonic functions as principal components:

$$\phi(x) = e^{ikx} \text{ with } \mu = \hat{C}_0(k) = |\hat{f}(k)|^2 \quad (2.4)$$

The conclusion is that the direct application of the method of principal components does not lead to a simple reconstruction of the signal.

Practical reconstruction

Practically speaking, the data of the wave are not available for all x and t and $C(x, y) \neq C_0(x - y)$. When the above method is then applied to the signal, the result will depend in an essential way on the length of these intervals.

By way of demonstration consider a wave profile of the form:

$$f(x) = \exp(-x^2/2)$$

with x in a spatial interval which is large enough to have no influences of the boundaries and $t \in [-T, T]$. For $T = 10, 20$ and 30 the first principal component ϕ_1 is shown in Fig 1.a. Its support increases with T ; for $T \rightarrow 0$, $\phi_1 = f$ and for $T \rightarrow \infty$ we get (2.4).

In Fig 1.b the number of modes necessary to cover a fraction of the "energy" (L^2 -norm) is shown for $T = 30$. The 95 % criterion indicates that the signal is rather complex, which is certainly not the case.

The conclusion is that a direct application of the method of principal components is *not suited* to detect the wave profile or velocity from a signal consisting of a translating wave at constant speed.

Modification

The performance of the POD method can be improved using the qualitative information that the signal represents some travelling wave. Considering the wave in a moving frame of reference separates the translation symmetry from the other (trivial) dynamics.

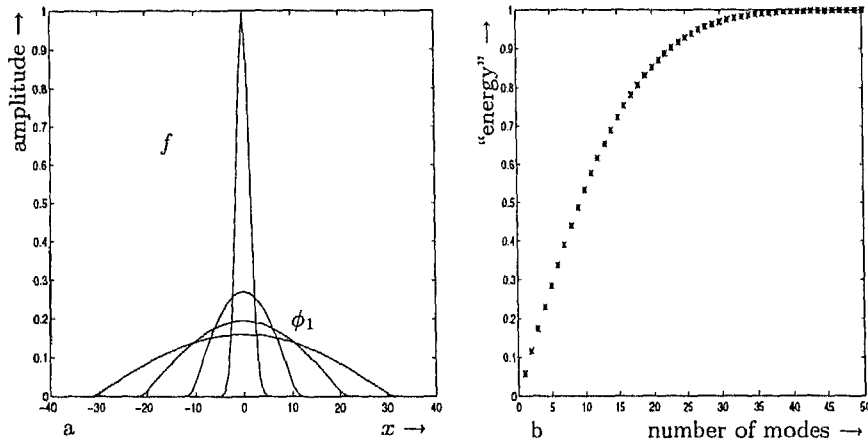


Figure 1. In a. we see the original signal f , and with an increasing support the principal topos ϕ_1 calculated from data for which $T=10,20$ and 30 respectively. In b. we see the number of modes necessary to cover a certain fraction of the "energy" of the signal when $T = 30$.

In a coordinate system moving with velocity λ the spatial autocorrelation function of the wave (2.3) reads:

$$C(x, y; \lambda) = \int f(x - (c - \lambda)t) f(y - (c - \lambda)t) dt \quad (2.5)$$

For $\lambda = c$ this reduces to $C(x, y; c) = T f(x) f(y)$ and the corresponding principal components read:

$$a_1 = T \|f\|, \phi_1 = f / \|f\|, \text{ and } a_i = 0 \quad \text{for } i \geq 2,$$

where T is the length of the time interval. For $\lambda \neq c$ we deal with a travelling wave with velocity $c - \lambda \neq 0$ and the POD method gives the same complications as described above.

A priori the velocity is not known and has to be determined from the signal too. To motivate the procedure, suppose the wave is known for all x and $t \in [-T, T]$ with $T < \infty$ and that it has compact support. There holds

$$\frac{1}{2T} \int_{-T}^T \left[\int u(x + \lambda t, t) \phi(x) dx \right]^2 dt \leq \int f^2(x) dx$$

for all λ and functions ϕ with $\int \phi^2 dx = 1$. Equality holds when $\lambda = c$ and $\phi = f(x) / \|f\|$, for $\lambda \neq c$ strict inequality holds. This shows that the velocity and the first principal component are characterized by:

$$\max_{\lambda, \phi} \left\{ \int \left[\int u(x + \lambda t, t) \phi(x) dx \right]^2 dt \mid \int \phi^2(x) dx = 1 \right\} \quad (2.6)$$

which is a generalization of (1.1) with (1.2) replaced by (2.5).

For a numerical treatment of (2.6) we have to discretise the signal; for example $u(x, t)$ becomes $u_{ij} = u(x_i, t_j)$ and ϕ becomes $\phi_i = \phi(x_i)$. The finite dimensional equivalent of (2.6) then reads

$$\max_{\lambda, \phi} \left\{ \underline{\phi}^* A(\lambda) \underline{\phi} \mid \underline{\phi}^* B \underline{\phi} \right\},$$

where $A(\lambda)$ is a matrix which components depend on λ and u_{ij} .

Solving (2.6) corresponds to the maximization of the eigenvalue $\mu(\lambda)$ of the generalized eigenvalue problem:

$$A(\lambda)\underline{\phi} = \mu(\lambda)B\underline{\phi}.$$

This largest eigenvalue can be obtained without calculation of $\underline{\phi}$ and smaller eigenvalues by use of the bisection or Given's method (Wilkinson *et al*, 1971).

Implementation of the above method has been done using piecewise linear basefunctions (IJzerman, 1996). The velocity and wave profile could be obtained up to machine precision when the grid size of the data is small enough.

3. Analysis of wave splitting

In this section we consider a substantially more difficult problem. The signal consists of numerical data of the following characteristic phenomenon in wave dynamics.

On a layer of water of constant depth, the theory of free surface waves predicts the existence of solitary waves: a single hump wave that is practically confined (exponentially decaying) and translates at a constant speed. Actually, a whole family of such waves exists: for every amplitude the solitary wave has a specific width and velocity:

$$u(x, t) = S(h, \lambda t, a)(x)$$

where S denotes the spatial profile of amplitude a on a layer of depth h and with its top at λt ; the speed λ depends on a . In a certain approximation (KdV) the profile function is given by

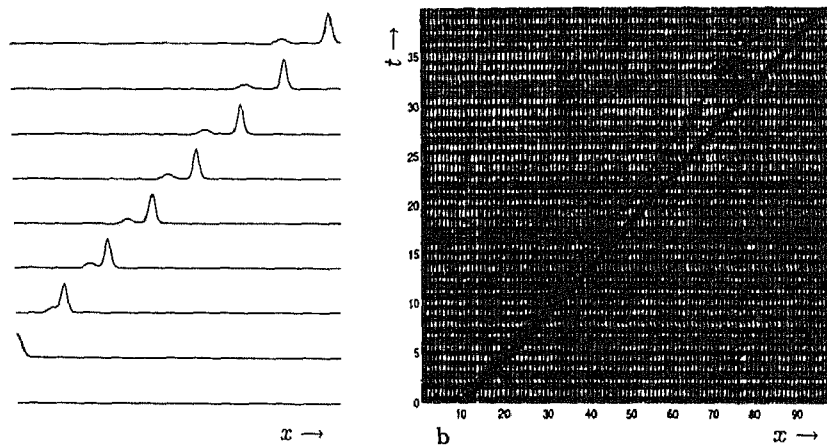
$$S(h, \phi, a) = a \operatorname{sech}^2[b(x - \phi)] \quad (3.7)$$

where the width b is related to a by (Van Groesen, 1996)

$$3a = 4b^2 h^3 \quad (3.8)$$

with an initial solitary wave above a flat bottom, the wave will when it runs into a region of different depth. In particular, when the depth decreases, nonlinear effects will become more dominant than linear effects, and the wave may undergo changes that resemble the splitting into several distinct waves. In an analytic approximation, it has been predicted that under certain circumstances a second wave will be emitted from the back of the first wave (Van Groesen, 1996; Van Groesen and Smit, 1993).

In this section we will verify this splitting process by analysing a signal consisting of numerical data. These data are obtained by calculating the evolution of a solitary wave during its run up from the exact water wave equations with a boundary element method, see (Van Daalen *et al*, 1996). The deformation of the wave is shown graphically; the appearance of a second wave is clearly visible.



Run up of a solitary wave showing the splitting into two waves. Results of numerical calculations; the initial amplitude of the wave is 10[cm], above a layer of depth 15[m]; the depth decreases linearly in the shaded area (15 to 22.5[m]) till 35[cm]; the computational domain is 100[m].

Using the data in more detail, a model manifold is taken that consists of the superposition of two solitary waves:

$$\mathcal{M} = \{ u(a_{+,-}, b_{+,-}, \phi_{+,-}) = a_{+} \operatorname{sech}^2[b_{+}(x - \phi_{+})] + a_{-} \operatorname{sech}^2[b_{-}(x - \phi_{-})] \mid a_{+,-}, b_{+,-}, \phi_{+,-} \} \quad (3.9)$$

where the parameters a, b, ϕ determine the amplitude, width and position of each wave respectively. Note that we do not prescribe the amplitude to be related by (3.8); see the remarks in the next section. The data are now used to determine the evolution of these parameters:

$$t \mapsto a_{+,-}(t), b_{+,-}(t), \phi_{+,-}(t).$$

The choice of a projection of the data on the manifold \mathcal{M} will determine this evolution, and hence the representation of the phenomenon. We use the simplest choice for this projection: at each instant, the L_2 -norm of the difference of the data with the manifold \mathcal{M} is minimized.

The results for the parameter evolution are shown in Fig 3; in Fig 4 the

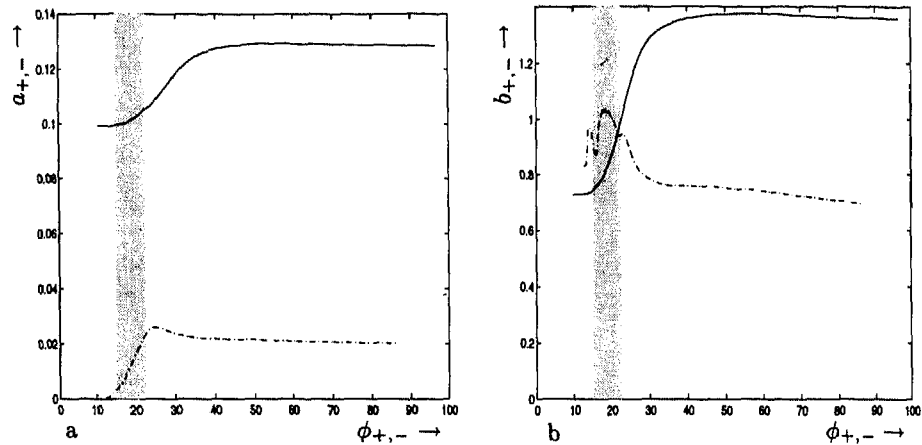


Figure 3. In a. we see the amplitude a_+ , (—) and a_- , (- - -) of the approximations within \mathcal{M} and in b. the width b_+ (—) and b_- (- - -) as a function of $\phi_{+,-}$.

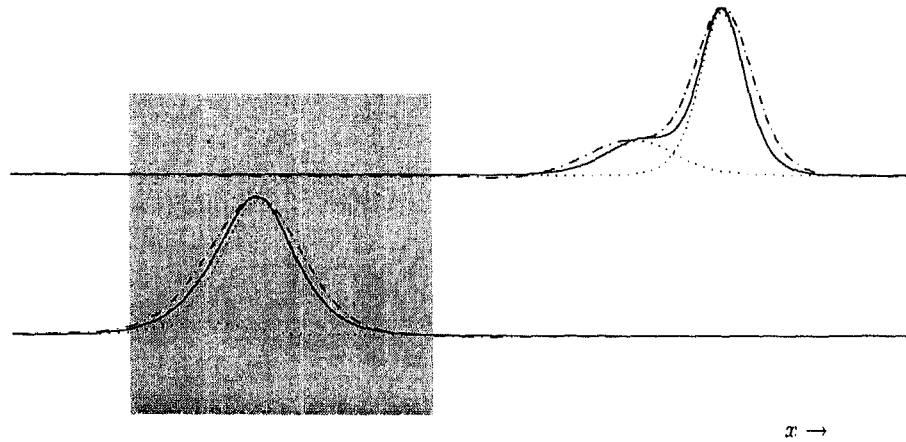


Figure 4. The approximation in the model manifold (—) at an instant during the splitting (at the left) and at an instant when the wave has passed the depth increase compared to the original data (- - -). The dotted lines are the constituent waves of the model manifold.

wave profile from the data is compared to the reconstructed element from \mathcal{M} at two characteristic times.

From these results we can draw the conclusion that the reconstruction using the manifold \mathcal{M} is quite successful: the splitting process is quantified in the

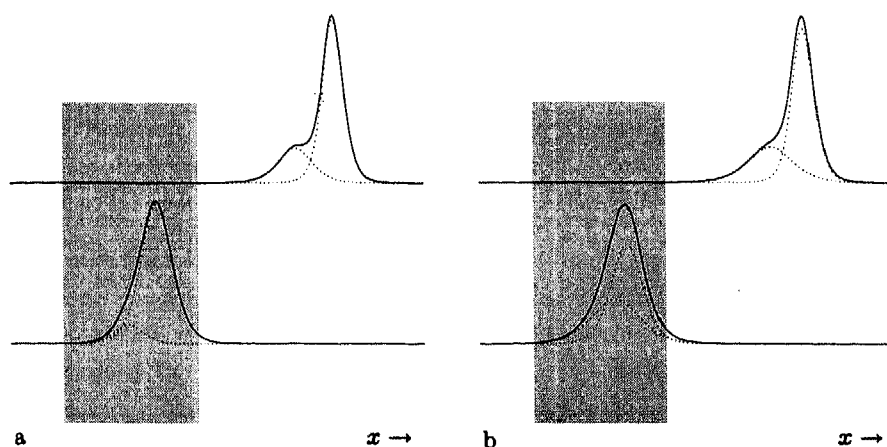


Figure 5. The approximation in the model manifold (—) at an instant during the splitting (at the left) and at an instant when the wave has passed the depth increase. In a. we see the result for the manifold \mathcal{M} and in b. for \mathcal{M}^* ; the dotted lines are the constituent waves.

evolution of the parameters of the manifold, and the reconstructed profile matches the profile of the data quite well. The evolution of width b_- as shown in Fig 3 is initially rather insecure; this can be expected since then this wave is still of small amplitude; the effect on the total wave profile is negligible. Furthermore, during the splitting process the assumption of the total profile as being a superposition of two waves is rather questionable; see also the next section.

4. Conclusions and remarks

In the previous section, the amplitude and width of the waves that determine the model manifold were taken as independent variables, and were determined independently in the minimization process. When the full free surface wave problem is approximated by a KdV-model, the solitary waves are of this form with amplitude and width related by (3.8). It is tempting to prescribe this relation in the definition of a smaller model manifold:

$$\mathcal{M}^* = \{ u(a_{+,-}, b_{+,-}, \phi_{+,-}) \in \mathcal{M} \mid a_{+,-} \text{ and } b_{+,-} \text{ related by (3.8)} \} \subset \mathcal{M}.$$

This implies that now only $a_{+,-}$ and the positions $\phi_{+,-}$ are the parameters to be adjusted from the signal. The projection of the signal on \mathcal{M}^* leads at first sight to different results: see Fig 6 where $b_{+,-}$ is calculated from $a_{+,-}$ according to relation (3.8). Observe that now the amplitude of the initial wave decreases drastically, while that of the new wave has much larger amplitude than previously. Remarkably enough, although the constituent waves that determine the profile by superposition are rather different for

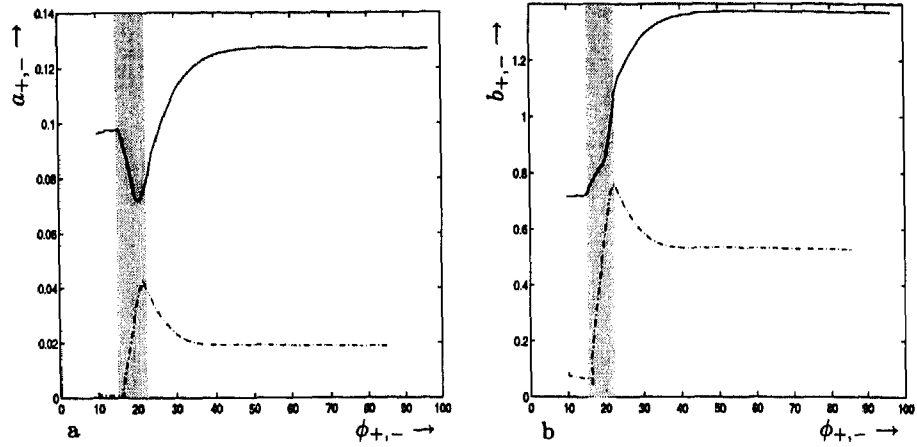


Figure 6. In a. we see the amplitude a_+ , (—) line and a_- , (- - -) line of the approximations within \mathcal{M}^* and in b. the width b_+ and b_- as a function of $\phi_{+,-}$.

the model manifold \mathcal{M} and \mathcal{M}^* , the actual superpositions agree quite well and both predict the profile of the signal quite well.

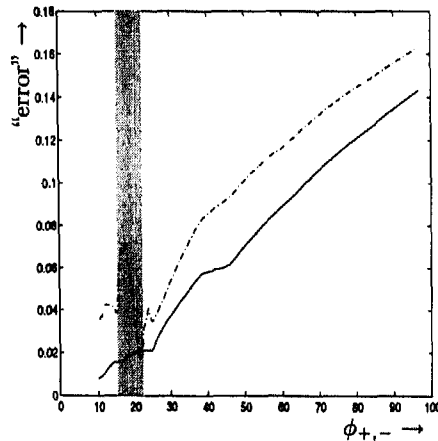


Figure 7. In this figure we see the relative "error" between the data and the approximation within \mathcal{M} , (—) and within \mathcal{M}^* , (- - -).

Fig 5 shows this complication that may arise in other, more complicated, cases as well. This picture shows that one and the same profile can be represented (in the order of approximation that is used) by two elements from the manifold \mathcal{M} that have two different constituent waves but give the same superposition:

$$u(a_{+,-}^1, b_{+,-}^1, \phi_{+,-}^1) = u(a_{+,-}^2, b_{+,-}^2, \phi_{+,-}^2).$$

The difference of the reconstruction in \mathcal{M} and that in \mathcal{M}^* should be a consequence of deviation of an actual solitary wave with the representation (3.7) where the parameters a, b are related to (3.8); freeing the parameters from this relation gives more flexibility ($\mathcal{M}^* \subset \mathcal{M}$) and hence a better approximation¹.

More quantitatively, the validity of the reconstruction can be investigated from Fig 7. The information in this picture should be interpreted with care. First the unexpected (almost linear) increase of the relative "error" has been found to be caused by errors in the numerical calculation of the waves, in particular by an incorrect treatment of the boundary conditions in the numerical grid which causes a non-zero level of the undisturbed water level. The remaining "error" in the reconstruction with the manifold \mathcal{M} can be attributed partly to an inaccurate description of the main waves with a sech^2 -profile, and partly to "radiation" that is known to emerge in a splitting process as considered here.

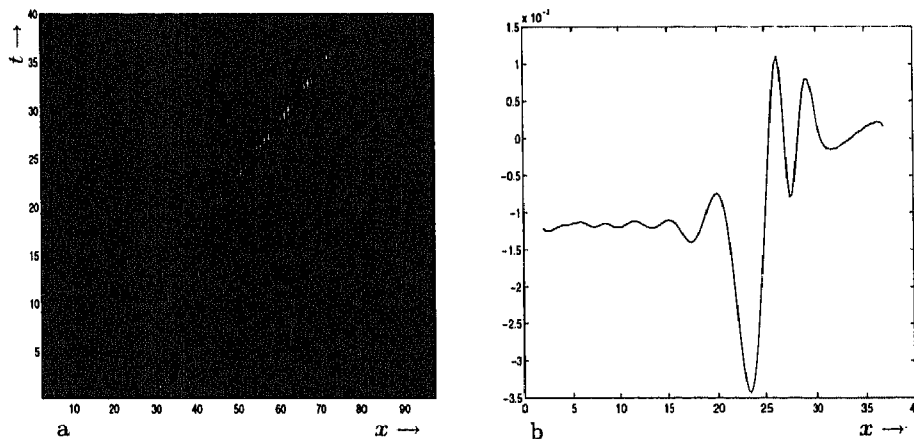


Figure 8. In a. we see the gray scale plot of the residual and in b. the first topos obtained in a moving frame of reference.

This "radiation" shows itself at the downstream side of the two solitary wave: there is a third wave with a maximum amplitude of ≈ 4 [mm], ($a_+ \approx 13$ [cm] and $a_- \approx 2$ [cm]). Since this wave is small it follows the linear wave equation and its velocity should be $c = \sqrt{gh} = 1.85 \dots$ [m/s].

To analyze this wave by use of the modified POD method we considered the residue: the signal minus the approximation in \mathcal{M} , see Fig 8.a. The white line from the lower left to the upper right is the third wave, the two

¹This is verified in a different way by analysing the data of numerical calculations of the wave problem in the approximation for which the solitary waves are exactly given by (3.7); then the reconstruction in \mathcal{M} and that in \mathcal{M}^* agree better.

dark lines on the right are the residues of the solitary waves. Also reflection is visible in this figure: when the solitary wave passes the uneven bottom, small waves are reflected to the left which are reflected once again at the boundary $x = 0$ of the computational domain. In the gray scale plot these waves are seen as the dark bands on the left of the third wave.

The first (normed) topois is shown in Fig 8.b; the oscillations on the right

Table 1. The estimated velocity from the POD method, for several temporal intervals: $[T, 40][s]$.

| T [s] | 30 | 25 | 20 | 15 | 10 | 5 |
|-----------------|-------|-------|-------|-------|-------|-------|
| \hat{c} [m/s] | 1.829 | 1.828 | 1.828 | 1.829 | 1.831 | 1.833 |

are the residues of the the larger solitary waves; the waves on the left are probably caused by dispersion.

The velocity obtained by the modified POD method is shown as a function of the temporal interval in Tab. 1. We see that the difference between the obtained and theoretical predicted velocity is within 2%, for all intervals used.

Concluding we can say that the proposed modification of the POD method is quite successful: the construction of the model manifold for the specific example of wave splitting allows one to determine the relevant parameters from the data; subtracting this primary phenomenon from the the data, the residue can be analyzed to obtain a more refined description of the contents of the data and a more refined model of the phenomenon.

References

- Aubry, N., Guyonnet, R. and Lima R. (1991) Spatiotemporal analysis of complex signals: Theory and applications. *Journal of Statistical Physics*, 64(3/4).
- Berkooz, G., Holmes, P. and Lumley, J.L. (1993) Th proper orthogonal decomposition in the analysis of turbulent flows, *Ann. Rev. Fluid Mech.*, volume 25.
- Van Daalen, E.F.G., Groesen, E., Pudjaprasetya, S.R. (1996) BEM-numerics and KdV-model analysis for solitary wave split-up. to appear in *Comp. Mech.*
- Van Groesen, E. (1996) A phenomenological description of soliton splitting during run up, to appear in *Contemp. Math.*
- Van Groesen, E. and Pudjaprasetya, S. R. (1993) Uni-directional Waves over slowly varying bottom. Part I: Derivation of a KdV-type of equation. *Wave Motion*, 18, 345-370.
- Sirovich, L. and Kirby, M. (1989) An eigenfunction approach to large scale transitional structures in jet flow. *Physical Fluids A*, 2(2):127-135.
- Wilkinson, J.H., and Reinsch, C. (1971) *Linear Algebra*, volume 2. Springer Verlag, New York.
- IJzerman, W.L. (1996) Model Manifolds and the reconstruction of waves, *Master Thesis*, University of Twente.