INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2000; **00**:1–6 *Prepared using* fldauth.cls *[Version: 2000/07/27 v1.0]*

The Monge-Kantorovitch mass transfer and its Computational Fluid Mechanics formulation

J.-D. Benamou *, Y. Brenier, K. Guittet

INRIA, Domaine de Voluceau, B.P.105 78153 Le Chesnay cedex, France.// Laboratoire d'analyse numérique, Université Paris 6, 4 place Jussieu, 75752 Paris cedex 05, France.

SUMMARY

This papers explains how Computational Fluid Mechanics (CFM) concepts can be used to solve Monge-Kantorovitch mass transfer type of problems(MKP). Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: Monge-Kantorovitch mass tranfer, Wasserstein distance, pressureless potential flow

1. INTRODUCTION

The interplay between the Monge-Kantorovitch mass transfer problem (MKP) and partial differential equation theory has recently experienced a surge of research activity and is now considered, in itself, as a branch of applied mathematics (see the surveys [12] and [16]).

In Fluid Dynamics, the MKP or equivalent concepts notably intervene in the semi-geostrophic equations [11] [10] [9] [3], an intermediate model in the geostrophic to primitive equation hierarchy used in meteorology. It is also used at the mathematical level in the existence theory of several Fluid Mechanics equations such as incompressible Euler with prescribed initial and final data [8].

From Monge's theory "des déblais et des remblais" in the late 18^{th} century to present days, a considerable amount of theoretical work has been gathered. The revival of this subject follows the pioneering work [6]. The (once again recent) interest of applied mathematicians in this problem has raised the question of the numerical resolution of the MKP in two and three dimensions (the resolution being trivial in 1-D). The efficiency of the classical Pogorelov construction (used in [11] and [15]) is difficult to evaluate. After and ad-hoc discretization the MKP can also be cast in a combinatorial optimization framework : the linear assignment problem. Even though there exists an, in principle, near optimal algorithm for this more general

*Correspondence to: INRIA, Domaine de Voluceau, B.P.105 78153 Le Chesnay cedex, France.

Contract/grant sponsor: Publishing Arts Research Council; contract/grant number: 98–1846389

Received
Revised

problem [1], the bridge from theory to practice turns out to be difficult to cross probably because of a huge unestimated constant affecting the cost of this simplex type method.

We recently established a "Computational Fluid Mechanics" (CFM) type formulation for the MKP in which the unknowns satisfy a simple gas dynamic model with prescribed initial and final densities [4]. Quite remarkably it is now the reverse application of CFM to the MKP which allows us to construct a robust and efficient numerical solver. Moreover, the CFM approach seems to be quite versatile. The formulation and the numerical method generalizes for instance to a multi-phasic context [5] and also to a mixed MKP/L^2 interpolation problem [2].

This paper reviews the material contained in [4], [5], [2] and [14].

2. THE MKP

A simple and modern formulation of the problem is the following : two bounded, positive measurable functions ρ_0 and ρ_T with compact support in $\mathbb{R}^d$ are given. They are called densities. We further require that they have the same mass, normalized to 1 :

$$\int_{\mathbb{R}^d} \rho_0(x)\,dx = \int_{\mathbb{R}^d} \rho_T(x)\,dx = 1. \tag{1}$$

The problem now is to select a mapping M from $\mathbb{R}^d$ to $\mathbb{R}^d$ which achieves the "transport" from ρ_0 to ρ_T in the following sense : for all Borel subsets $A \subset \mathbb{R}^d$, M satisfies

$$\int_{M^{-1}(A)} \rho_0(x)\,dx = \int_A \rho_T(x)\,dx. \tag{2}$$

Equation (2) is a weak formulation of the so called Jacobian equation

$$det(DM(x))\rho_T(M(x)) = \rho_0(x), \tag{3}$$

which can be derived when M is a smooth one-to-one map. Here $det(DM)$ is the determinant of the Jacobi matrix of M and represents the rate of compression or spreading of the mass induced by the map $x \mapsto M(x)$ (we recall that ρ_0 and ρ_T and M must satisfy (3)). The Jacobian problem is clearly under-determined (when transferring Dirac masses for instance, any permutation of the loaded points gives rise to an admissible mapping) and it is natural to select among the maps satisfying (2) those which are optimal in a suitable sense. The Monge-Kantorovitch problem consists in choosing the mapping M which satisfies the constraint (2) and minimizes the "transportation" cost

$$C(M) = \int_{\mathbb{R}^d} \|x - M(x)\|^2 \rho_0(x) dx. \tag{4}$$

Roughly speaking, $\|x-M(x)\|^2\rho_0(x)$ is the travelled distance squared, weighted by the amount of the transferred mass. This point-wise interpretation is restrictive and relies on the hypothesis that M is smooth and one-to-one. The class of mappings M satisfying (2) is of course much wider and allows for instance to pointwise "split" or "coalescing" mass. Considering this wider class is also the key to the main theoretical results [6].

Theorem 2.1. *There is a unique optimal mapping $\bar{M}$ defined on the support of ρ_0 satisfying (2). The mapping M is characterized as the unique mapping from this class which can be written as the gradient of a convex potential Φ :*

$$\bar{M}(x) = \nabla\Phi(x). \tag{5}$$

The optimal value of the cost (4) also is a distance (squared) called Wasserstein distance between densities ρ_0 and ρ_T. This distance is usually defined as :

$$d_{Wa}(\rho_0, \rho_T)^2 = \inf_{\mu} \int |x-y|^2 d\mu(x,y), \tag{6}$$

where μ spans the space of probability measures $\mathbb{R}^d \times \mathbb{R}^d$ with marginals ρ_0 and ρ_T. We therefore have

$$C(\bar{M}) = d_{Wa}(\rho_0, \rho_T)^2. \tag{7}$$

Regarding the applications of the MKP in Fluid Mechanics, we again refer the reader to the papers, (excellent) surveys and monographs cited in the Introduction.

3. A CFM FORMULATION FOR THE MKP

The MKP problem described in the above section is a quadratic space minimization problem in M, with a non-linear, non-convex and highly degenerate constraint (2). We do not know how to numerically enforce this constraint so we instead proposed a reformulation of the problem based on the introduction of a "time" variable t. We first describe the CFM type problem and then give its relationship with the classical MKP.

We fix a time interval $[0,T]$ and consider all possible smooth enough, time-dependent, density and velocity fields, $\rho(t,x) \geq 0$, $v(t,x) \in \mathbb{R}^d$, subject to the continuity equation

$$\partial_t \rho + \nabla.(\rho v) = 0 \tag{8}$$

for $0 < t < T$ and $x \in \mathbb{R}^d$, and the initial and final conditions

$$\rho(0,.) = \rho_0, \quad \rho(T,.) = \rho_T. \tag{9}$$

Then, our new problem is the minimization of the action

$$K(\rho, v) = T \int_{\mathbb{R}^d} \int_0^T \rho(t,x)|v(t,x)|^2 dx dt, \tag{10}$$

amongst all (ρ, v) satisfying (8) and (9).

Let us point out that a Continuum Mechanics formulation was already implicitly contained in the original problem addressed by Monge : "le problème des remblais et des déblais". Eliminating the time variable was just a clever way of reducing the dimension of the problem. However, from a computational point of view, reintroducing the time variable allows to solve a convex (although not quadratic) space-time minimization problem in the density and momentum variables, namely ρ and $m = \rho v$, with linear constraints. Indeed under this change

of variable, the integrand of (10) is $\frac{|m(t,x)|^2}{2\rho(t,x)}$ which can be written as the Legendre transform of the indicatrix function of the convex set $K = \{\{a,b\} : \mathbb{R}\times\mathbb{R}^d \to \mathbb{R}\times\mathbb{R}^d, \ \ s.\,t. \ \ a + \frac{|b|^2}{2} \leq 0 \ \textit{pointwise}\}$

$$\frac{|m(t,x)|^2}{2\rho(t,x)} = \sup_{\{a,b\}\in K} [a(t,x)\rho(t,x) + b(t,x).m(t,x)] \tag{11}$$

and $K(\rho, v)$ is convex in $(\rho, m = \rho v)$. Finally constraint (9) is unchanged but (8) becomes linear

$$\partial_t \rho + \nabla.m = 0. \tag{12}$$

This is, in our opinion, a considerable advantage, in spite of the addition of the extra (but not artificial) time variable. In addition, the Continuum Mechanics formulation, provides a natural time interpolant $\rho(t,x)$ of the data ρ_0 and ρ_T and a velocity field $v(t,x)$ which moves ρ_0 toward ρ_T.

The relation between the classical MKP and this "CFM" MKP, formally established in [4], is given by the following proposition

Proposition 3.1. *The square of the Wasserstein distance is equal to the infimum of*

$$T\int_{\mathbb{R}^d}\int_0^T \rho(t,x)|v(t,x)|^2 dxdt,$$

among all (ρ, v) satisfying (8) and (9).

There exists moreover a unique optimal flow

$$X(0,x) = x, \ \ \partial_t X(t,x) = v(t, X(t,x))$$

given in terms of the potential Φ

$$X(t,x) = x + \frac{t}{T}(\nabla\Phi(x) - x).$$

A rigorous Hilbertian framework for this problem can be found in [14].

We finally remark that the optimality conditions of this space-time minimization problem are

$$v(t,x) = \nabla\phi(t,x), \tag{13}$$

where the potential ϕ is the Lagrange multiplier associated with the constraints (8), (9), and the Hamilton-Jacobi equation

$$\partial_t\phi + \frac{1}{2}|\nabla\phi|^2 = 0. \tag{14}$$

In terms of Fluid Mechanics, it means that the optimal solution of (14) (8) is given by a pressureless potential flow.

4. TWO GENERALIZATIONS OF THE MKP

4.1. The multi-phasic mass transfer problem

We here consider a problem which involves multiple phases (ρ_α, v_α) $(\alpha = 1, M)$ with the constraint that the total mass is prescribed by a given function $\bar{\rho}$

$$\sum_{\alpha=1}^{M} \rho_\alpha = \bar{\rho} \tag{15}$$

Each phase individually satisfies (8)

$$\partial_t \rho_\alpha + \nabla(\rho_\alpha v_\alpha) = 0, \tag{16}$$

initial and final conditions like (9)

$$\rho_\alpha(0,.) = \rho_\alpha^0, \ \rho_\alpha(T,.) = \rho_\alpha^T \tag{17}$$

where the ρ_α^0 and ρ_α^1 are given nonnegative functions which satisfy the compatibility conditions

$$\int \rho_\alpha^0(x)\,dx = \int \rho_\alpha^T(x)\,dx\,, \alpha = 1,..,M. \tag{18}$$

and

$$\int \sum_{\alpha=1}^{M} \rho_\alpha^0(x)\,dx = \int \bar{\rho}(t,x)\,dx\,, t \in [0,T]. \tag{19}$$

The generalization of the MKP now is to minimize the sum of the kinetic energies

$$K(\rho, v) = \frac{1}{2} \sum_{\alpha=1}^{M} \int_0^T \int \rho_\alpha(t,x)|v_\alpha(t,x)|^2 dxdt \tag{20}$$

where $\rho = (\rho_1, ..., \rho_M)$ and $v = (v_1, ..., v_M)$ satisfy the constraints (15-17) above.

This multi-phasic "CFM" MKP problem can be solved numerically using the same technique as in the mono-phasic case (see [5]). The optimality conditions for the Lagrange multipliers (ϕ_α) associated to the constraints (16) and a new multiplier p for the global constraint (16) are

$$\partial_t \phi_\alpha + \frac{|\nabla_x \phi_\alpha|^2}{2} + \frac{p}{\mu_\alpha} = 0\,, \alpha = 1,..,M. \tag{21}$$

These equations are again very similar to the optimality equations for the mono-phasic problem. However, because of the incompressibility constraint (15) a pressure p appears in the Hamilton-Jacobi equation (21), inducing the expected coupling between all phases. Notice that, taking $\bar{\rho} \equiv 1$, we recover the homogenized vortex-sheet model discussed in [7] where the existence of "variational solutions" for this problem is proven.

4.2. A mixed MKP/L^2 interpolation problem

Before explaining this mixed problem, let us point out that it is possible to express the L^2 distance using a similar time-dependent formulation. We consider the minimization problem

$$d^2_{L^2}(\rho_0,\rho_1) = \inf_{\rho,v}\{\frac{1}{T}\int_{\mathbb{R}^d}\int_0^T |\partial_t\rho(t,x)|^2 dxdt\}, \tag{22}$$

with (ρ, v) again subject to the constraints (8) and (9). The optimization problem does not depend on v anymore and the constraint (8) is just mentioned here by analogy with section 2.1. We again have a convex minimization problem. The cost function appearing on the right hand side of (22) can be differentiated, and the optimality conditions simply express the fact that the optimal $\rho(t,x)$ satisfies (9) and

$$\partial^2_{tt}\rho = 0. \tag{23}$$

The optimal solution is therefore directly given by the time interpolation formula

$$\rho(t,x) = \frac{\rho_T(x)-\rho_0(x)}{T}t + \rho_0(x). \tag{24}$$

Replacing $\partial_t\rho$ in (22), we obtain the claimed L^2 distance.

So it is quite tempting to make an intrinsic interpolation between the L^2 and Wasserstein distances. The mixed distance is defined likewise by

$$d_{was/L^2}(\rho_0,\rho_1)^2 = \inf_{\rho,v}\{T\int_{\mathbb{R}^d}\int_0^T \theta\rho(t,x)|v(t,x)|^2 + (1-\theta)\frac{|\partial_t\rho(t,x)|^2}{2}dxdt\} \tag{25}$$

where (ρ, v) must again satisfy the constraints (8) and (9). The parameter $\theta \in [0,1]$ is the interpolation parameter. Of course as $\theta = 1$ or $\theta = 0$ we recover respectively the Wasserstein or the L^2 distances.

A discussion about the motivation for this object and an illustration of the rather different properties of the L^2 and Wasserstein distances can be found in [4] where a numerical method is proposed to solve (25).

5. NUMERICAL RESOLUTION

The L^2 MKP can be written as a saddle-point problem by introducing a space-time dependent Lagrange multiplier $\phi(t,x)$ for constraints (8) and (9). The Lagrangian is given by :

$$L(\phi,\rho,m) = \int_0^T\int_D [\frac{|m|^2}{2\rho} - \partial_t\phi\rho - \nabla_x\phi.m] - \int_D[\phi(0,.)\rho_0 - \phi(T,.)\rho_T], \tag{26}$$

where the terms involving ϕ come from (8) by integration by part and using the boundary conditions (9).

Given initial and final densities ρ_0 and ρ_T, the L^2 MKP is equivalent to the saddle-point problem :

$$\inf_{\rho,m}\sup_{\phi} L(\phi,\rho,m). \tag{27}$$

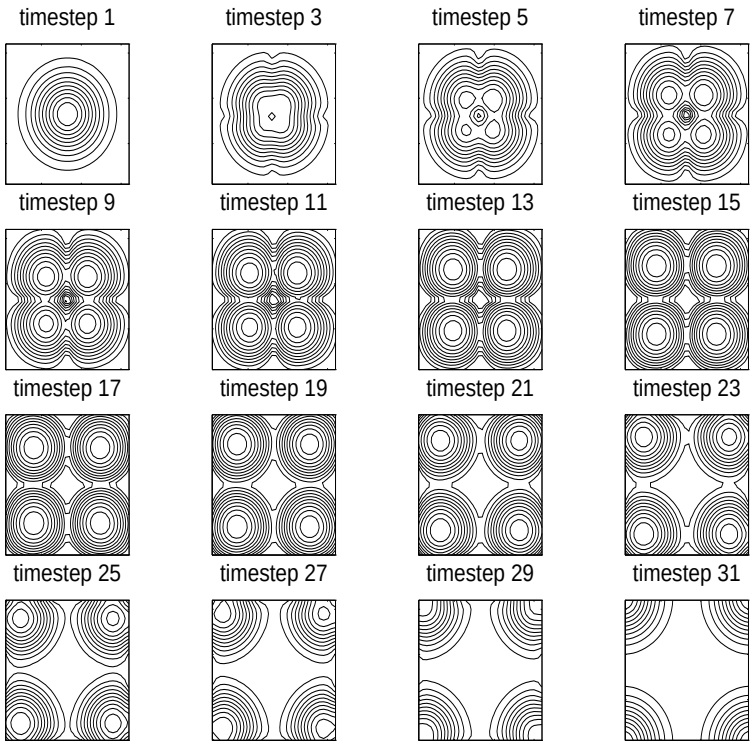

Figure 1. Contours plots of the density at successive time steps

This saddle-point formulation can be recast (using (11)) in a form matching the problem of the elastoplastic deformation of a cylindrical rod as presented in [13]. In this note, the Lagrangian is first "augmented" and an efficient algorithm from [13], called ALG2, based on relaxations of the Uzawa algorithm is used. We follow the same line as in [4] (and refer to this paper for more details) to solve the problem. We get a three step iterative method which constructs a sequence converging to the saddle-point. The more expensive part of the algorithm is the iteration of a space-time Laplace equation.

Our favorite test case is periodic in space and we therefore implicitly consider an infinite grid of similar (here square) cells. The initial density is an infinite array of Gaussian functions centered inside the cells and the final density is the same array but the Gaussian functions have been shifted to the corners of the cells. As discussed in [4] the optimal Wasserstein mapping splits the Gaussian in four and sends each part to the corners. The optimal transfer is not a simple translation, as one would superficially think, but is rather one that splits each Gaussian function into four pieces, sending each of them to the nearest corner (Figure 1).

A second approach consists in the elimination of the "two point boundary value problem" by relaxing the final constraint on the density $\rho(T, x) = \rho_T$ and enforce it using a penalization

term added to the cost function. We have used this approach to treat the general mixed MKP/L^2 problem to which the Augmented Lagrangian technique does not seem to apply. Thus we consider

$$d_{relax}(\rho_0,\rho_1)^2 = \inf_m \{ T \int_{\mathbb{R}^d} \int_0^T \theta \frac{|m|^2}{2\rho} + (1-\theta)\alpha \frac{|\partial_t \rho|^2}{2} dx dt + \int_{\mathbb{R}^d} \gamma \frac{|\rho(T,x) - \rho_T(x)|^2}{2} dx, \} \tag{28}$$

where γ is a positive penalty parameter. Note that ρ is no more a minimization variable but a state variable solution of (8) where m and the initial conditions $\rho(0,x) = \rho_0$ are given (we have dropped the final condition). This approach simplifies considerably this optimization problem which now becomes an optimal control problem. A classical technique to evaluate the gradient of the cost function is to use direct/adjoint problems. We have done so and embedded this approach into a conjugate gradient algorithm.

We now present the computation for different values of θ. The test case is similar to the one presented for the augmented Lagrangian method but a small constant has been added to the density. Indeed this second algorithm is very sensitive to the absence of mass. This feature is not shared by the Augmented Lagrangian method which seems robust when dealing with zero values for the mass (e.g., we can treat the case involving characteristic functions).

Figure 2 shows the contours plots of the densities at successive time steps for three different values of $\theta : \theta = 1$ is the pure Wasserstein problem which splits the Gaussian in four parts and translates the mass to the corners, $\theta = 0$ is the pure L^2 problem which is a simple pointwise in space and linear in time interpolation between initial and final density and $\theta = 0.5$ is a mixed problem which "interpolates" between the two extreme behavior. For all computations $\gamma = 10$.

Finally, we show a simulation of a bi-phasic transport. The global mass $\bar{\rho}$ is a constant. The initial and final density for the first phase is similar to our first test case (figure 1) and the second phase is completely determined by (15). The numerical method arises from the augmented Lagrangian technique [5].

The level lines of the first density are given in figure 3. The global mass $\bar{\rho}$ is chosen small enough so that the mass cannot concentrate above a given level and this effect competes with the transport of the split four parts of the Gaussian observed in Figure 1. The mass is tranferred in an elongated almond shape to lower its maximum level.

REFERENCES

1. M. L. Balinski. A competitive (dual) simplex method for the assignment problem. *Math. Programming* 1986; **34**(2):125–141.
2. J.-D. Benamou, Y. Brenier. Mixed l^2/wasserstein optimal mapping between prescribed density functions. *J. Optim. Theory Appl. , to appear* 2001; **111**(2):225–271.
3. J.-D. Benamou, Y. Brenier. Weak existence for the semi-geostrophic equations formulated as a coupled monge-ampère/transport problem. *SIAM J. Applied Math.* 1998; **58**(5):1450–1461.
4. J.-D. Benamou, Y. Brenier. A computational fluid mechanics solution to the monge-kantorovich mass transfer problem. *Numer. Math.* 2000; **84**:375–393.
5. J.-D. Benamou, Y. Brenier, K. Guittet. Numerical resolution of a multiphasic optimal mass transport problem. *preprint*.

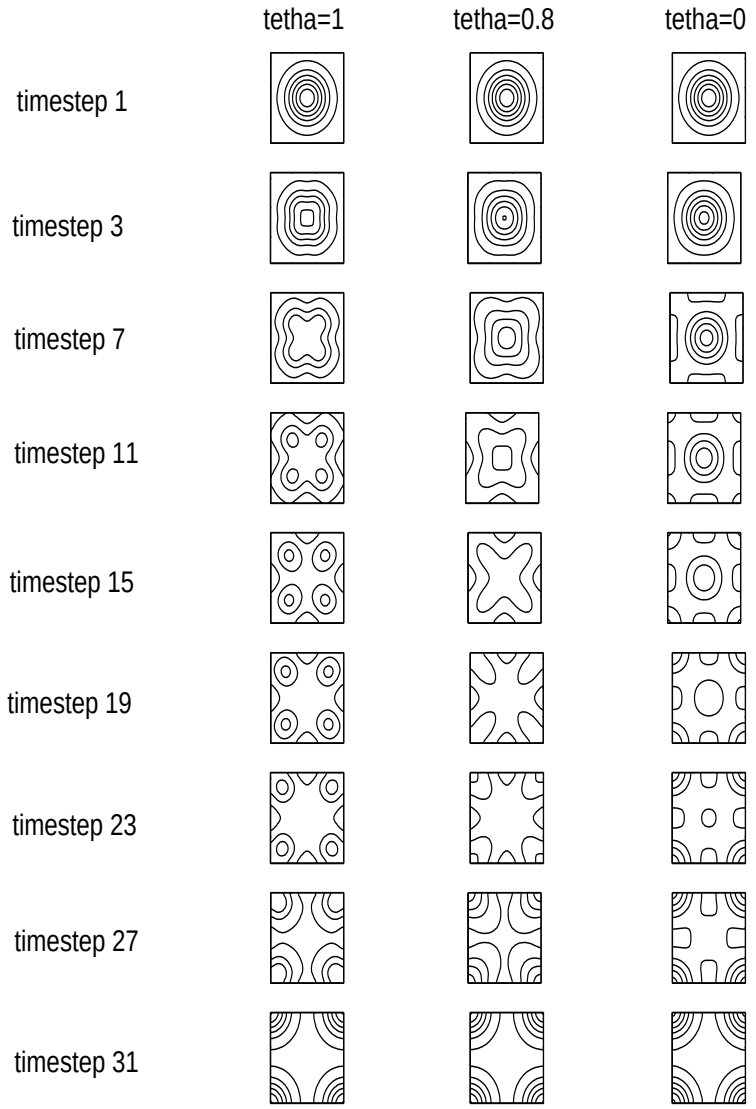

Figure 2. Contours plots of the density at successive time steps for the Conjugate Gradient algorithm, $\beta = 1, 0.8, 0,\ \gamma = 10$.

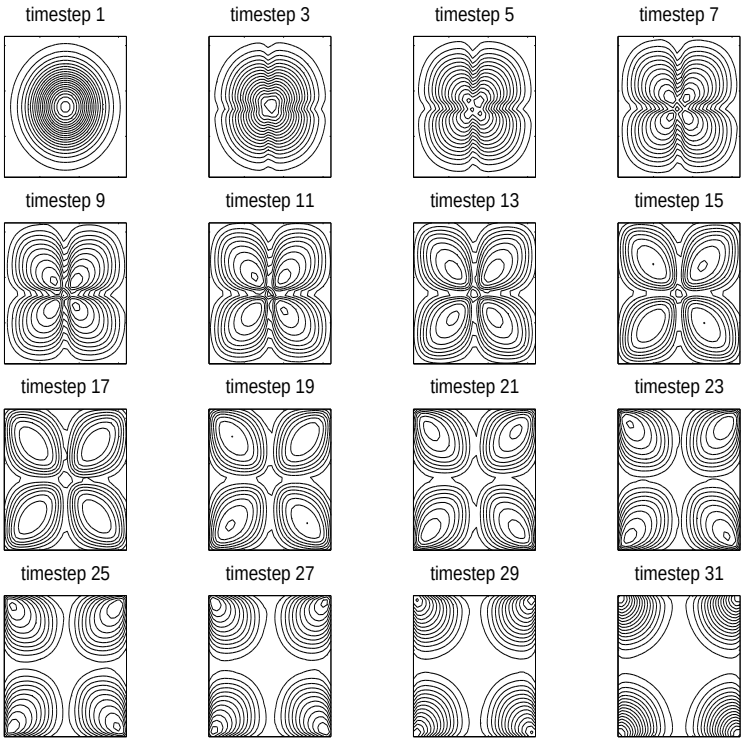

Figure 3. Contours plots of the density at successive time steps for the first phase (8)

6. Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.* 1991; **44**(4):375–417.
7. Yann Brenier. *A homogenized model for vortex sheets* Arch. Rational Mech. Anal., 138(4):319–353, 1997.
8. Yann Brenier. Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations. *Comm. Pure Appl. Math.* 1999; **52**(4):411–452.
9. M. J. P. Cullen, W. Gangbo. A variational approach for the 2-d semi-geostrophic shallow water equations. *preprint*.
10. M. J. P. Cullen, J. Norbury, R. J. Purser. Generalised Lagrangian solutions for atmospheric and oceanic flows. *SIAM J. Appl. Math.* 1991; **51**(1):20–31.
11. M. J. P. Cullen, R. J. Purser. An extended Lagrangian theory of semigeostrophic frontogenesis. *J. Atmospheric Sci.* 1984; **41**(9):1477–1497.
12. L.C. Evans. *Partial differential equations and Monge-Kantorovich mass transfer* Lecture Notes, 1998.
13. Michel Fortin, Roland Glowinski. *Augmented Lagrangian methods* North-Holland Publishing Co., Amsterdam, 1983.
14. K. Guittet. An hilbertian framework for the time-continuous monge-kantorovitch problem. *preprint*.
15. Sergey A. Kochengin, Vladimir I. Oliker. Determination of reflector surfaces from near-field scattering data. II. Numerical solution *Numer. Math.* 1998; **79**(4):553–568.
16. C. Villani, F. Otto, Y. Brenier. *In preparation*.