

## A MODIFIED LEAST ACTION PRINCIPLE ALLOWING MASS CONCENTRATIONS FOR THE EARLY UNIVERSE RECONSTRUCTION PROBLEM

YANN BRENIER

*CNRS, FR 2800, Université de Nice Sophia,  
Parc Valrose, FR06108 Nice, France  
brenier@math.unice.fr*

Received 20 December 2010

*In Memory of Michelle Schatzman*

We address the early universe reconstruction (EUR) problem (as considered by Frisch and coauthors in [26]), and the related Zeldovich approximate model [46]. By substituting the fully nonlinear Monge–Ampère equation for the linear Poisson equation to model gravitation, we introduce a modified mathematical model (“Monge–Ampère gravitation/MAG”), for which the Zeldovich approximation becomes exact. The MAG model enjoys a least action principle in which we can input mass concentration effects in a canonical way, based on the theory of gradient flows with convex potentials and somewhat related to the concept of self-dual Lagrangians developed by Ghoussoub [29]. A fully discrete algorithm is introduced for the EUR problem in one space dimension.

### 0. Introduction

This paper addresses the early universe reconstruction (EUR) problem discussed by Frisch and coauthors in [26, 18], following Peebles’ seminal paper [38]. In these references, gravitation is not modeled according to the full Einstein equations, but rather to a semi-Newtonian approximation, where classical Newtonian interactions just take place in an Einstein–de Sitter space, corresponding to a big bang scenario. In suitable coordinates, the model can be described as follows. Let us denote, for each gravitating body, its label by  $a$  and its position at time  $t$  by  $X(t, a) \in \mathbb{R}^3$ . The density field  $\rho$  is defined by

$$\rho(t, x) = \int_a \delta(x - X(t, a)) \quad (0.1)$$

and the gravitational potential  $\varphi(t, x)$  satisfies

$$\rho = 1 + t\Delta\varphi. \quad (0.2)$$

The Newton law for each gravitating bodies is just

$$\partial_t(\alpha(t)^2 \partial_t X(t, a)) = -t^{-1} \beta(t)^2 (\nabla \varphi)(t, X(t, a)), \tag{0.3}$$

where  $\alpha$  and  $\beta$  are time-dependent scaling parameters provided by general relativity (GR). Following [26, 18] (case of an Einstein–de Sitter universe), we set:

$$\alpha(t) = t^{3/4}, \quad \beta(t) = t^{3/4} \sqrt{3/2}. \tag{0.4}$$

In the case of coefficients (0.4), we find

$$\frac{2t}{3} \partial_{tt}^2 X(t, a) + \partial_t X(t, a) = -(\nabla \varphi)(t, X(t, a)). \tag{0.5}$$

Notice that, in this model, which we call SNS (as semi-Newtonian system), friction dominates at early times. (In some sense, Newton modified by Einstein returns to Aristoteles.) Remarkably enough, at  $t = 0$ , the density field must be uniformly equal to 1 (otherwise solutions get unbounded) and the velocity is enslaved by the gravity potential term. Thus, we can write

$$\rho(0, x) = 1, \quad X(0, a) = a, \quad \partial_t X(0, a) = -\nabla \varphi_0(a), \quad \Delta \varphi_0 = \lim_{t \downarrow 0} \frac{\rho(t, x) - 1}{t}. \tag{0.6}$$

So, at time  $t = 0$ , the gravitational matter behaves as a continuum, with a definite (and potential) velocity field. Consistently with the SNS, such a continuum may keep, at least for a while, a potential velocity field  $v = v(t, x) = \nabla \theta(t, x)$  such that

$$\partial_t X(t, a) = v(t, X(t, a)),$$

for all labels  $a$ . Then, Newton’s law (0.3) can be expressed in terms of  $\theta$  and  $\varphi$  as:

$$\partial_t(\alpha^2 \theta) + \alpha^2 \frac{|\nabla \theta|^2}{2} + t^{-1} \beta^2 \varphi = 0, \quad \partial_t X(t, a) = (\nabla \theta)(t, X(t, a)). \tag{0.7}$$

There is no room for a discrete repartition of gravitational matter at this early stage and only the time evolution is able to progressively produce discrete structures such as isolated particles (or, more generally, concentrated matter on sheets or filaments), as the density field  $\rho$  becomes singular with respect to the Lebesgue measure. As a matter of fact, the SNS (0.1), (0.2), (0.7) may (and usually does) produce collisions in finite time, as will be seen later on, which generate such concentrations. Another remarkable feature of the SNS is that, at time  $t = 0$ , the only possible initial condition is the density fluctuation field  $\rho'_0$  (or, equivalently, the initial gravitational potential  $\varphi_0$ ) defined by

$$\rho'_0(x) = \lim_{t \downarrow 0} \frac{\rho(t, x) - 1}{t} = \Delta \varphi_0(x). \tag{0.8}$$

This fluctuation field is of paramount importance in the study of the *very* early universe [30], which is of great interest in high energy physics and quantum gravity theory. Since the evolution in time of the model depends only on  $\rho'_0$ , it is plausible that one could recover this field from the simple observation of a comparable scalar

field at our present time, say  $t = T$ . A natural candidate is obviously the present density field  $\rho_T = \rho(t = T, \cdot)$ . This is precisely the early universe reconstruction (EUR) problem. To solve the EUR, a remarkable fact can be used. The SNS enjoys a least action principle, with a strictly convex action! (This is a very exceptional situation, in physics and mechanics, where action principles generally correspond to saddle points of functionals of indefinite type [22].) More precisely:

**Proposition 0.1.** *Any smooth solution  $(\rho, v = \nabla\theta, \varphi)$  of the SNS (0.1), (0.2), (0.7), on some time interval  $0 < t_0 < t < t_1$ , is characterized by the following least action principle: as  $\rho$  is fixed at time  $t = t_0$  and  $t = t_1$ ,  $(\rho, \rho v)$  is the unique minimizer (with respect to compactly supported perturbations) of the strictly CONVEX action*

$$\int_{t_0}^{t_1} dt \int (\alpha(t)^2 \rho(t, x) |v(t, x)|^2 + \beta(t)^2 |\nabla\varphi(t, x)|^2) dx, \tag{0.9}$$

under the linear constraints

$$\partial_t \rho + \nabla \cdot (\rho v) = 0, \quad t \Delta \varphi = \rho - 1. \tag{0.10}$$

This result extends to the limit case  $t_0 = 0, t_1 = T$ , with  $\rho$  given at time  $T$  and  $\rho(t = 0, \cdot) = 1$ . Thus, the reconstruction of the early universe looks easy: knowing  $\rho_T$  we just have to minimize a strictly convex action and we recover the whole solution of the SNS for  $t \in [0, T]$ ! (Mathematically speaking, this problem has been addressed by Loeper in [33].) Unfortunately, this reasoning does not take into account that smooth solutions to the SNS may break down in finite time, due to the concentration of the density field which may become singular with respect to the Lebesgue measure. The goal of this paper is to investigate how the action can be modified so that its minimizers are not necessarily concentration-free. A similar problem, in the framework of adhesion-fragmentation processes, has been recently solved by Wolansky [45]. (See also the pioneering work of Shnirelman [41] for sticky particles and adhesion dynamics.) Our approach is different and more reminiscent of the recent theory of self-dual Lagrangians by Ghoussoub [29]. Unfortunately, our method does not apply to the desired SNS, but rather to the modified system obtained by substituting the fully nonlinear Monge–Ampère equation  $\rho = \det(I + tD_x^2\varphi)$ , for the Poisson equation  $\rho = 1 + t\Delta\varphi$ . We call this new model “Monge–Ampère gravitation” (MAG). There is no difference between SNS and MAG for solutions depending only on one spatial coordinate (i.e. with sheet structure) and they are formally asymptotically close for  $t \rightarrow 0$ . Of course, changing the model is not a satisfactory approach, without further justification. Our main argument is the following remarkable property of the MAG system: it admits as exact solutions some approximate solutions to the SNS, suggested by Zeldovich [46] (and (1.11) below). As a secondary justification, let us recall that the SNS is, after all, itself an approximation of the full Einstein equations and it might be, from this viewpoint, equally good to use the Monge–Ampère equation and the Poisson

equation. (A similar situation occurs in fluid mechanics when comparing the quasi-geostrophic and the semi-geostrophic approximations of the Euler equations for ocean and atmosphere dynamics, as discussed in [23]. See also [19].) However, there will be no attempt in this paper to justify this last statement.

The structure of the paper is as follows: In Sec. 1, we review Zeldovich' approximation to the SNS. Then, in Sec. 2, we introduce the MAG action and the corresponding MAG equations.

In Sec. 3, we observe that the potential part of the MAG action has the very special property to be a squared distance function. This allows a rewriting of the action as an exact square and we find as special minimizers all the solutions of the gradient flow equation associated to the potential, with, among them, all the Zeldovich solutions. (These special solutions play more or less the same role as "instantons" in Yang–Mills theory [29].) It turns out that this gradient flow belongs to a very well-studied class of evolution equations with "maximal monotone operators" [20]. This suggests a somewhat canonical modification of the action.

In Sec. 4, we introduce a fully discrete algorithm for the numerical minimization of the MAG action.

In Sec. 5, we introduce a numerical scheme for the initial value problem and, finally, in Sec. 6, we provide numerical results in the very special case of one space variable.

## 1. Zeldovich Approximations to the SNS

An amazingly simple approximate formula was proposed for solutions of the SNS (0.1), (0.2), (0.7) by Zeldovich [46]:

$$X(t, a) = a - t\nabla\varphi_0(a), \quad (1.11)$$

with

$$\rho(0, x) = 1, \quad \Delta\varphi_0 = \rho'_0(x) = \lim_{t \downarrow 0} \frac{\rho(t, x) - 1}{t}.$$

This formula turns out to be exact for small time and initial conditions depending only on one space coordinate (this will be seen below). The Zeldovich approximate formula predicts mass concentrations in finite time. Indeed, denoting by  $\Lambda$  the largest eigenvalue of the Hessian matrix  $D^2\varphi_0(a)$ , for all  $a$ , we see that, whenever  $\Lambda > 0$ , the map  $a \rightarrow X(t, a)$  is no longer invertible at  $t = \Lambda^{-1}$ . Beyond the concentration time, there are many possibilities of extending the formula and this is still a controversial issue from the physical viewpoint. It depends very much on whether or not we want to prevent interpenetration of particles. If we do so, we are naturally led to the model of adhesion dynamics, where particles merge after collisions, which is the most possible dissipative behavior beyond concentrations. (See [40, 43, 8, 5, 25, 42, 17, 39].) This issue can simply be addressed in terms of nonlinear hyperbolic PDEs [24]. Indeed, given a Zeldovich solution  $X$  defined by

(1.11), let us introduce the field  $u(t, x)$  implicitly defined by:

$$u(t, X(t, a)) = \frac{a - X(t, a)}{t} = \nabla\varphi_0(a), \tag{1.12}$$

as long as  $a \rightarrow X(t, a)$  stays smooth and invertible. Then, we see that  $u$  solves the multidimensional “inviscid Burgers” equation

$$\partial_t u + (u \cdot \nabla)u = 0. \tag{1.13}$$

In one space dimension, if we want a global solution for all times, the monotonicity condition  $\partial_a X(t, a) \geq 0$  exactly corresponds to “Oleinik’s entropy condition”  $\partial_x u \leq 1/t$ , which guarantees both global existence and uniqueness for solutions of the inviscid Burgers equation (1.13), written in “conservation form”

$$\partial_t u + \partial_x \left( \frac{u^2}{2} \right) = 0. \tag{1.14}$$

## 2. Monge–Ampère Gravitation

### 2.1. An abstract framework for Monge–Ampère gravitation

Let  $H$  be a (separable) Hilbert space  $H$  equipped with its norm denoted  $\| \cdot \|$  and the corresponding inner product  $(\cdot, \cdot)$ . We first consider the general dynamical system

$$\frac{d^2 X}{dt^2} = (\nabla_H \Phi)[X], \tag{2.15}$$

where  $t \rightarrow X(t)$  is valued in  $H$ ,  $\nabla_H$  denotes the gradient operator in  $H$ , and  $\Phi$  is a given “potential” defined on  $H$ . (Observe that we do not follow the usual sign convention for the potential, for notational convenience.) As is well known, such a system admits a least action principle, at least at a formal level. Indeed, for a curve  $t \rightarrow X(t)$  valued in the Hilbert space  $H$ , we may define its action between times  $t_0$  and  $t_1$ ,  $t_1 > t_0$  by:

$$A_{[t_0, t_1]}[X] = \int_{t_0}^{t_1} \frac{1}{2} \left\| \frac{dX}{dt} \right\|^2 + \Phi[X(t)] dt. \tag{2.16}$$

Then, the dynamical equation (2.15) can be seen as the formal optimality equation obtained by minimizing the action (2.16) as the endpoints  $X(t_0)$  and  $X(t_1)$  are fixed.

Next, we crucially assume the potential to be of the form:

$$\Phi[X] = \inf \left\{ \frac{\|X - s\|^2}{2}; s \in S \right\}, \tag{2.17}$$

where  $S$  is a given bounded subset of  $H$ . Then, when it makes sense,  $X - \nabla_H \Phi[X]$  is just the closest point  $\pi[X]$  to  $X$  in the set  $S$ . (Clearly this definition is ambiguous whenever  $X$  has several distinct closest points, which may happen unless  $S$  is a

convex set. In some cases,  $X$  may have no closest point in  $S$ !) As a consequence, (2.15) formally means:

$$\frac{d^2 X}{dt^2} = X - \pi[X], \tag{2.18}$$

where  $\pi[X]$  is the closest point to  $X$  on  $S$ . With this formulation, we can guess a large class of explicit solutions. Indeed, let us assume that  $X(0) = X_0$  has a unique closest point  $\pi[X_0] = \pi_0$  on  $S$ . Then the linear (but not convex) combination of  $X_0$  and  $\pi_0$  given by:

$$X(t) = \pi_0 + e^t(X_0 - \pi_0) \tag{2.19}$$

solves (2.18) as long as  $\pi_0$  stays the unique projection of  $X(t)$ . Intuitively,  $X(t)$  gets repelled from its initial position in the opposite direction of its closest point on  $S$ , keeping for a while  $\pi_0$  as its closest point on  $S$  until a new point in  $S$  gets even closer. Whenever  $S$  is a convex set, this repulsion mechanism provides an obvious global solution. Indeed, all points contained in the infinite segment  $\{\pi_0 + r(X_0 - \pi_0), r \geq 0\}$  admits  $\pi_0$  as their unique closest point on  $S$ . In the case of a non-convex set  $S$ , this is not true in general and formula (2.19) is able to provide no more than a local solution. The situation is very clear in the elementary case when  $S$  is the unit sphere in  $H$ . Then, 0 is the unique point where  $\Phi$  is not differentiable. We get as special solution

$$X(t) = r_0^{-1}(1 + (r_0 - 1)e^t)X_0,$$

where  $X_0 \neq 0$  and  $r_0 = \|X_0\|$ . We see that, if  $r_0 < 1$ , then the solution reaches 0 at time  $T = -\log(1 - r_0)$  and its continuation beyond  $T$  gets ambiguous.

**Miscellaneous mathematical remarks.** (1) The potential  $\Phi$  given by (2.17) is a smooth perturbation of a Lipschitz concave function; indeed:

$$\Phi[X] = \frac{\|X\|^2}{2} - \Pi[X], \tag{2.20}$$

where  $\Pi$  is the Lipschitz convex functional defined by:

$$\Pi[X] = \sup \left\{ ((X, s)) - \frac{\|s\|^2}{2}; s \in S \right\}. \tag{2.21}$$

A classical result of convex analysis [4] asserts that, for every Lipschitz convex function defined on a Hilbert space, the set  $\tilde{H}$  where the function is differentiable is always “fat” in the topological sense of Baire: namely  $\tilde{H}$  is dense and contains a countable intersection of dense open subsets of  $H$  [4]. In the particular case of  $\Pi$ , the set of differentiability  $\tilde{H}$  is contained in the set of all points  $X$  in  $H$  for which there is a unique closest point  $s = \pi[X]$  on  $S$ . Thus, the potential  $\Phi$  defined by (2.17) is everywhere differentiable on  $\tilde{H}$  and its gradient in  $H$  is given by:

$$\nabla_H \Phi[X] = X - \pi[X], \quad \forall X \in \tilde{H}. \tag{2.22}$$

(2) In the case when  $H$  is the finite-dimensional Hilbert space  $\mathbb{R}^n$ , for such a potential (namely a smooth perturbation of a concave Lipschitz function), the dynamical system (2.15) has a unique global solution for Lebesgue almost every initial condition  $(X(0), X'(0)) \in \mathbb{R}^{2n}$  and is, therefore, well-posed in the sense of Bouchut and Ambrosio [9, 1]. To the best of our knowledge, there is no similar theory in infinite dimension and the well-posedness of (2.15) is then a challenging open. (A somewhat related attempt is the theory developed by Ambrosio and Gangbo for some infinite dimensional Hamiltonian systems [2]. See also [19, 27, 28].)

### 2.2. Monge–Ampère gravitation

**Definition.** Since the dimension 3 does not matter in the definition of the MAG model, we consider a smooth bounded closed domain  $D \subset \mathbb{R}^d$ . We assume  $D$  to be of unit Lebesgue measure. The MAG model is defined by choosing for  $H$  the Hilbert space of all Lebesgue square-integrable maps from  $D$  to  $\mathbb{R}^d$ ,

$$H = L^2(D, \mathbb{R}^d), \tag{2.23}$$

and for  $S$  the subset of all Lebesgue measure-preserving maps  $s$  of  $D$ :

$$S = \left\{ s \in H, \int_D f(s(a))da = \int_D f(a)da, \forall f \in C(\mathbb{R}^d) \right\}. \tag{2.24}$$

In addition, with respect to the abstract framework, we input coefficients  $\alpha, \beta$  given by (0.4) and substitute

$$\int_{t_0}^{t_1} \frac{1}{2} \alpha^2(t) \left\| \frac{dX}{dt} \right\|^2 + t^{-2} \beta^2(t) \Phi[X(t)] dt \tag{2.25}$$

for (2.16) and get as optimality equations:

$$t^2 \beta^{-2}(t) \frac{d}{dt} \left( \alpha^2(t) \frac{dX}{dt} \right) = (\nabla_H \Phi)[X] = X - \pi[X]. \tag{2.26}$$

Using tools of optimal transport theory (see the Appendix), the right-hand side of this equation can be more concretely written

$$X(t, a) - \pi[X(t, \cdot)](a) = -t \nabla \varphi(t, X(t, a)), \tag{2.27}$$

where  $\varphi = \varphi(t, x)$  solves a Monge–Ampère equation

$$\det(I + t D_x^2 \varphi(t, x)) = \rho(t, x), \tag{2.28}$$

where  $\rho$  is the density field

$$\rho(t, x) = \int_D \delta(x - X(t, a)) da. \tag{2.29}$$

Thus, we have obtained the MAG system with

$$t\beta^{-2}(t)\frac{d}{dt}\left(\alpha^2(t)\frac{dX}{dt}\right) = -(\nabla\varphi)(t, X(t, a)). \tag{2.30}$$

### 3. A Modified Action Taking into Account Concentrations

#### 3.1. Modified action in the abstract framework

In this section, we go back to the abstract framework of a potential  $\Phi$  defined as the squared distance to a bounded subset  $S$  of a general Hilbert space  $H$ , according to (2.18), (2.20), (2.21). Since potential  $\Phi$  is a squared distance to some subset  $S$  inside  $H$ , it solves, at least formally, the stationary Hamilton–Jacobi equation:

$$\Phi = \frac{\|\nabla_H\Phi\|^2}{2}, \tag{3.31}$$

where  $\nabla_H$  denotes the gradient operator in  $H$ . This suggests to rewrite, at least formally, the action (2.25) as

$$\begin{aligned} &\int_{t_0}^{t_1} \frac{1}{2} \left\| \frac{dX}{dt} \right\|^2 + \frac{\|\nabla_H\Phi[X(t)]\|^2}{2} dt \\ &= \int_{t_0}^{t_1} \frac{1}{2} \left\| \frac{dX}{dt} - \nabla_H\Phi[X(t)] \right\|^2 + \left( \left( \frac{dX}{dt}, \nabla_H\Phi[X(t)] \right) \right) dt \\ &= \Phi[X(t_1)] - \Phi[X(t_0)] + \int_{t_0}^{t_1} \frac{1}{2} \left\| \frac{dX}{dt} - \nabla_H\Phi[X(t)] \right\|^2 dt. \end{aligned} \tag{3.32}$$

Under this “self-dual” form (see [29] for a systematic study of “self-dual Lagrangians”), it is obvious that any solution of

$$\frac{dX}{dt} = (\nabla_H\Phi)[X] = X - (\nabla_H\Pi)[X] \tag{3.33}$$

is always a minimizer of the action as  $X(t_0)$  and  $X(t_1)$  are fixed (just like instantons in Euclidean Yang–Mills theory, cf. [29]).

As already mentioned, in spite of the rather nice structure (2.20) of  $\Phi$ , as a quadratic perturbation of a convex Lipschitz function, the corresponding second-order equation (2.18) is not so well understood. In sharp contrast, the first-order equation (3.33) is a standard “gradient flow” equation (GF), that can be solved by classical “maximal monotone operator” theory [20].

In the framework of maximal monotone operator theory, Eq. (3.33) is usually written as a sub-differential inclusion:

$$-\frac{dX}{dt} + X \in \partial\Pi[X], \tag{3.34}$$



which is well-posed in  $H$  since  $\Pi$  is Lipschitz and convex. Here, we use standard notations of convex analysis, for which  $\partial$  denotes the sub-differential of a convex function [20]:

$$\partial\Pi[X] = \{Z \in H; \Pi[Y] \geq \Pi[X] + ((Z, Y - X)), \forall Y \in H\}. \tag{3.35}$$

A remarkable property [20] of each solution  $X(t) \in H$  is to be not only a Lipschitz continuous function of  $t$  but also right-differentiable at each  $t$  with

$$-\frac{dX(t+0)}{dt} + X(t) = d^0\Pi[X(t)], \quad \forall t, \tag{3.36}$$

where  $d^0\Pi[X]$ , following [3], denotes the element of  $\partial\Pi[X]$  with minimal norm (which is uniquely defined):

$$\|d^0\Pi[X]\| = \min\{\|s\|; s \in \partial\Pi[X]\}. \tag{3.37}$$

Finally, notice that  $X(t)$  is a locally Lipschitz function of  $t$  with values in the separable Hilbert space  $H$ ,  $X(t)$  is therefore almost everywhere differentiable in  $t$  by Rademacher theorem. Since  $X(t)$  is right-differentiable everywhere, we conclude that:

$$-\frac{dX}{dt} + X(t) = d^0\Pi[X(t)], \tag{3.38}$$

holds true both in the almost everywhere sense and in the sense of distributions.

Our main point is now to introduce a *modified* action. There are two possible ways to do it. First, we may introduce the modified potential  $\tilde{\Phi}$ :

$$\tilde{\Phi}[X] = \frac{1}{2}\|X - d^0\Pi[X]\|^2 \tag{3.39}$$

and the corresponding modified action

$$\begin{aligned} \tilde{A}_{[t_0, t_1]}[X] &= \int_{t_0}^{t_1} \frac{1}{2} \left\| \frac{dX}{dt} \right\|^2 + \tilde{\Phi}[X(t)] dt \\ &= \int_{t_0}^{t_1} \frac{1}{2} \left\| \frac{dX}{dt} \right\|^2 + \frac{1}{2} \|X - d^0\Pi[X]\|^2 dt. \end{aligned} \tag{3.40}$$

Alternately, sticking more closely to the self-dual formulation, we may directly modify the action by setting

$$\hat{A}_{[t_0, t_1]}[X] = \int_{t_0}^{t_1} \frac{1}{2} \left\| \frac{dX}{dt} - X + d^0\Pi[X] \right\|^2 dt. \tag{3.41}$$

It is not clear to us that these modified actions coincide (up to boundary terms). Nevertheless, we will take the second option, mostly for numerical purposes, because it leads to simpler algorithms.

**3.2. Modified action for the MAG model**

We now consider the MAG model. This means, with respect to the abstract framework, that  $H$  and  $S$  are now defined by (2.23), (2.24) and (2.30) substitutes for (2.15).

In order to take into account coefficients  $(\alpha, \beta)$  (given by (0.4)), we first rewrite the action as:

$$A = \int_{t_0}^{t_1} \alpha(t)^2 \left\| \frac{dX}{dt} \right\|^2 + t^{-2} \beta(t)^2 \|\nabla_H \Phi[X(t)]\|^2 dt. \tag{3.42}$$

As in the homogeneous case  $\alpha = \beta = 1$ , we keep in mind that

$$\frac{1}{2} \|\nabla_H \Phi[X(t)]\|^2 = \Phi[X(t)]$$

and look at the cross-term:

$$J = \int_{t_0}^{t_1} \alpha(t)t^{-1}\beta(t) \left( \left( \frac{dX}{dt}, \nabla_H \Phi[X(t)] \right) \right) dt = \int_{t_0}^{t_1} \alpha(t)t^{-1}\beta(t) \frac{d}{dt} (\Phi[X(t)]) dt.$$

By integration by part, we get

$$\begin{aligned} J & - \alpha(t_1)t_1^{-1}\beta(t_1)\Phi[X(t_1)] + \alpha(t_0)t_0^{-1}\beta(t_0)\Phi[X(t_0)] \\ & = - \int_{t_0}^{t_1} \Phi[X(t)] \frac{d}{dt} (\alpha(t)t^{-1}\beta(t)) dt \\ & = - \frac{1}{2} \int_{t_0}^{t_1} \|\nabla_H \Phi[X(t)]\|^2 \frac{d}{dt} (\alpha(t)t^{-1}\beta(t)) dt \\ & = - \frac{\lambda}{2} \int_{t_0}^{t_1} t^{-2}\beta(t)^2 \|\nabla_H \Phi[X(t)]\|^2, \end{aligned}$$

provided we assume

$$\frac{d}{dt} (\alpha(t)t^{-1}\beta(t)) = \lambda t^{-2}\beta(t)^2, \tag{3.43}$$

for some constant  $\lambda$ , which is consistent with data (0.4) if we choose  $\lambda = 1/\sqrt{6}$ . From this calculation of the cross-term  $J$ , we deduce that the action  $A$  defined by (3.42) can be written:

$$A = BT + \int_{t_0}^{t_1} \left\| \alpha(t) \frac{dX}{dt} - \mu t^{-1} \beta(t) \nabla_H \Phi[X(t)] \right\|^2 dt,$$

where  $BT$  is a boundary term depending only on  $X(t_1)$  and  $X(t_0)$ , provided  $\mu^2 + \mu\lambda = 1$ . For data (0.4), we get  $\lambda = 1/\sqrt{6}$  and  $\mu = \sqrt{2/3}$ . Therefore, all solutions of the gradient-flow equation

$$\alpha(t) \frac{dX}{dt} = \mu t^{-1} \beta(t) \nabla_H \Phi[X(t)] \tag{3.44}$$

automatically are minimizers of the action (3.42). For data (0.4), this gradient-flow equation reduces to:

$$t \frac{dX}{dt} = \nabla_H \Phi[X(t)] = X(t) - \nabla_H \Pi[X(t)]. \tag{3.45}$$

The gradient-flow equation should be understood in the more precise sense:

$$t \frac{dX(t+0)}{dt} = X(t) - d^0 \Pi[X(t)], \tag{3.46}$$

which takes concentration into account, globally in time. In some sense, formulation (3.46) not only allows concentrations but guarantees the largest possible dissipation of kinetic energy during the concentration process (which is of course questionable from the physical viewpoint.) Accordingly, we suggest, for the MAG model, the following modified action:

$$\hat{A} = \int_{t_0}^{t_1} t^{-1/2} \left\| t \frac{dX}{dt} - X + d^0 \Pi[X] \right\|^2 dt. \tag{3.47}$$

### 3.3. Zeldovich solutions

Special solutions of (3.45) can be obtained, thanks to the concept of “rearrangements with convex potential” as follows. By definition, the MAG model relies on the set  $S$  of all Lebesgue measure-preserving maps (2.24). This set contains the identity map  $\text{Id}$  as an obvious element. The set  $K \subset H$  of all points  $X$  which admits  $\text{Id}$  as a closest point on  $S$  plays a crucial role. It can be characterized (cf. Appendix on optimal transportation theory), as the convex cone of all maps  $X \in H$  with a convex potential, which means that there is a convex function  $\psi$  defined on  $\mathbb{R}^d$  and valued in  $] -\infty, +\infty ]$  which is almost everywhere differentiable on  $D$  with  $\nabla \psi(a) = X(a)$ , a.e. on  $D$ . It turns out that any map  $X \in H$  has a unique rearrangement  $X^\sharp$  in  $K$ , which means

$$\int_D \delta(x - X^\sharp(a)) da = \int_D \delta(x - X(a)) da$$

(cf. the Appendix).

Therefore, special solutions of (3.45) can be obtained, by looking for solutions  $X(t)$  valued in the convex cone  $K$  of all maps with convex potential. Indeed, for such solutions, we have:

$$\nabla_H \Phi[X(t)] = X(t) - \nabla_H \Pi[X(t)] = X(t) - \pi[X(t)] = X(t) - \text{Id}$$

and (3.45) reduces to the linear ODE

$$t \frac{dX}{dt} = X - \text{Id} \tag{3.48}$$

as long as  $X(t)$  belongs to  $K$ , i.e.  $X(t, a) = \nabla\psi(t, a)$ , with  $\psi(t, a)$  convex in  $a$ . This leads to the explicit formula:

$$X(t, a) = \nabla\psi(t, a) = a + \frac{t}{t_0}(\nabla\psi(t_0, a) - a) = a + \frac{t}{t_0}(X(t_0, a) - a), \quad (3.49)$$

as long as  $\psi$  stays convex in  $a$ . This exactly coincides with Zeldovich formula (1.11) discussed in the Introduction. Remarkably enough, for Monge–Ampère gravitation, Zeldovich approximation (1.11) is just exact!

### 3.4. Modified action in one space dimension

Let us focus on the one space dimension case when:  $D = [-1/2, 1/2]$ . Then, the modified potential  $\tilde{\Phi}$  can be explicitly computed in the case of a piecewise smooth map  $Y$  valued in  $K$ . Indeed, in one space dimension, maps in  $K$ , with convex potential are just increasing maps. So, there is a finite number of plateaux  $[a_j, b_j]$  on which  $Y$  is constant with values  $Y_j$  and outside of which  $Y$  is a piecewise smooth strictly increasing function.

Notice that the corresponding image-measure  $\rho(dx)$  defined by

$$\rho(dx) = \int_D \delta(x - Y(a))da$$

has a singular part  $\rho_s$  given by:

$$\rho_s(dx) = \sum_j (b_j - a_j)\delta(x - Y_j).$$

Then  $d^0\Pi[Y]$  (the element of the sub-differential  $\partial\Pi[Y]$  with minimal  $L^2$  norm) coincides with the identity map outside of the plateaux and takes value  $(a_j + b_j)/2$  inside  $[a_j, b_j]$ .

After elementary calculations, we find

$$\begin{aligned} \|Y - d^0\Pi[Y]\|^2 &= \|Y\|^2 - 2((Y, \text{Id})) + \|\text{Id}\|^2 - \sum_j \int_{a_j}^{b_j} \left(a - \frac{a_j + b_j}{2}\right)^2 da \\ &= \|Y - \text{Id}\|^2 - \frac{1}{12} \sum_j (b_j - a_j)^3. \end{aligned}$$

Here we very clearly see the discrepancy between the original potential  $\Phi$  and the modified potential  $\tilde{\Phi}$ :

$$\tilde{\Phi}[Y] = \Phi[Y] - \frac{1}{24} \sum_j (b_j - a_j)^3. \quad (3.50)$$

**Remark 1.** Specialists of nonlinear hyperbolic conservation laws will recognize in the second term of this expression the very expression of the so-called “entropy production” term for the inviscid Burgers equation (1.14), written in material coordinates [7, 24].

### 3.5. Eulerian version of the gradient flow equation

The gradient flow equation (3.44) has a Eulerian version. Indeed, the corresponding measures  $(\rho, \rho v)$ , defined by

$$\rho(t, x) = \int_D \delta(x - X(t, a)) da, \quad \rho v(t, x) = \int_D \partial_t X(t, a) \delta(x - X(t, a)) da, \quad (3.51)$$

are (formal) solutions of the following system of PDE:

$$\partial_t \rho - \nabla \cdot (\rho \nabla \varphi) = 0, \quad \rho = \det(I + t D_x^2 \varphi). \quad (3.52)$$

This model can be seen as a fully nonlinear counterpart of various models popular in biology (chemotaxis) or astronomy, involving the Poisson equation — or other linear equations involving a singular Green function — rather than the Monge–Ampère equation. A common feature of all these models is their ability at describing concentration phenomena [32, 31, 36, 21].

## 4. A Discrete Action for the MAG Model

### 4.1. A time-discrete scheme for the gradient flow equation

In view of numerical calculations, our first step is to get a time-discrete version of the modified action. Instead of directly getting a discrete version of (3.47), it seems wiser to us to start from a time-discrete version of the gradient flow equation (3.45).

A natural candidate is:

$$X_{n+1} = X_n + (X_n - \nabla_H \Pi[X_n]) \theta_n + \eta_n, \quad (4.53)$$

where  $X_n$  is an approximation of  $X(t)$  at the  $n$ th time-step  $T_n$ , for  $n = 0, \dots, N$ ,  $T_0 = t_0$ ,  $T_N = t_1$ , with  $\theta_n = \frac{T_{n+1}}{T_n} - 1 \downarrow 0$ . (In the special case  $t_0 = 0$ , it is natural to set  $X_0 = \text{Id}$  and to provide  $X_1$  as the initial condition.) In formula (4.53),  $\eta_n$  is a small perturbation added to the discrete solution so that, for every  $n$ ,  $X_n$  is a point of differentiability of  $\Pi$ . Thus,  $\nabla_H \Pi[X_n]$  is well-defined and is also the closest point  $\pi[X_n]$  to  $X_n$  in  $S$ . Indeed, as a smooth perturbation of a Lipschitz concave function on  $H$ ,  $\Pi$  is differentiable on a “fat” dense subset  $\tilde{H}$  of  $H$  (i.e. containing a countable intersection of dense open sets). Thus, we may choose a perturbation term  $\eta_n$ , arbitrarily small, so that  $X_n$  falls in the “good” set  $\tilde{H}$  where  $\nabla_H \Pi$  is well-defined. By doing so, we do not generate a big error. Indeed, we keep control on the cumulated error, thanks to the following stability estimate for two distinct solutions  $X_n, \tilde{X}_n$  of (4.53) (where we neglect the perturbation terms  $\eta_n, \tilde{\eta}_n$  for notational simplicity),

$$\|X_{n+1} - \tilde{X}_{n+1}\|^2 \leq (1 + \theta_n)^2 \|X_n - \tilde{X}_n\|^2 + c_0 \theta_n^2, \quad (4.54)$$

where  $c_0$  is the squared diameter of  $S$ . [Indeed, we get from (4.53)

$$\begin{aligned} & \|X_{n+1} - \tilde{X}_{n+1}\|^2 \\ &= (1 + \theta_n)^2 \|X_n - \tilde{X}_n\|^2 - 2(1 + \theta_n)\theta_n ((X_n - \tilde{X}_n, \nabla_H \Pi[X_n] - \nabla_H \Pi[\tilde{X}_n])) \\ & \quad + \theta_n^2 \|\pi[X_n] - \pi[\tilde{X}_n]\|^2 \end{aligned}$$

and observe that the second term on the right-hand side is less than zero since  $\Pi$  is convex, and the third one is dominated by  $c_0\theta_n^2$ .]

As a matter of fact, this stability estimate is also essentially sufficient to prove the convergence of the scheme as  $\theta_n \downarrow 0$  to the continuous model (3.46), for the uniform convergence in time with respect to the strong topology of  $H$ . (See [14, 15] for examples of similar results for various nonlinear hyperbolic conservation laws.) Notice that concentration phenomena, which are present at the continuous level, are correctly taken into account by the time-discrete scheme, in spite of the fact that the discrete scheme never involves the computation of  $d^0\Pi$ , which is a big advantage in practice!

**4.2. A time-discrete action for the MAG model**

From the time-discrete scheme (4.53) for the gradient-flow equation, we define a time discrete version of modified action (3.47) just by setting:

$$\sum_{n=0}^{N-1} r_n \|X_{n+1} - X_n - (X_n - \pi[X_n])\theta_n\|^2, \tag{4.55}$$

with  $r_n = \frac{T_n^{3/2}}{T_{n+1} - T_n}$ . Nevertheless, in view of the EUR problem, it is more reasonable to minimize the time-discrete action (4.55) when the data are not  $X_0$  and  $X_N$  but rather the corresponding probability measures  $\rho_0$  and  $\rho_N$  defined by:

$$\rho_0(dx) = \int_D \delta(x - X_0(a))da, \quad \rho_N(dx) = \int_D \delta(x - X_N(a))da.$$

So there is a big loss of information (since the same probability measure can be generated by a continuum of maps). This problem can be addressed in terms of rearrangements with convex potentials. As a matter of fact, fixing  $\rho_0$  and  $\rho_N$  is equivalent to fixing the rearrangements with convex potentials  $X_0^\sharp$  and  $X_N^\sharp$ , rather than  $X_0$  and  $X_N$  themselves. It is very fortunate that, one can rewrite the discrete scheme (4.53) as a self-consistent scheme for the rearrangement  $Y_n = X_n^\sharp$  with convex potential. Indeed, let us assume, for simplicity, that, at each  $n$ , the solution of the scheme  $X_n$  has a polar factorization  $X_n = Y_n \circ s_n$  (cf. the Appendix), where  $Y_n = X_n^\sharp \in K$  is the unique rearrangement with convex potential of  $X_n$  and  $s_n = \pi[X_n] \in S$  is the closest point in  $S$  to  $X_n$ . Then, we can rewrite (4.53) as:

$$Y_{n+1} \circ s_{n+1} = (Y_n + (Y_n - \text{Id})\theta_n) \circ s_n.$$

But, this implies that  $Y_{n+1}$  is the unique rearrangement of  $Y_n + (Y_n - \text{Id})\theta_n$  with a convex potential. In other words, we have a well-defined self-consistent scheme for  $Y_n \in K$ , namely:

$$Y_{n+1} = (Y_n + (Y_n - \text{Id})\theta_n)^\sharp. \tag{4.56}$$

Accordingly, the minimization of the time-discrete MAG action can be seen, in “polar coordinates”  $(Y_n, s_n) \in K \times S$ , as the minimization of

$$\sum_{n=0}^{N-1} r_n \|Y_{n+1} \circ s_{n+1} - (Y_n + (Y_n - \text{Id})\theta_n) \circ s_n\|^2, \tag{4.57}$$

with  $r_n = \frac{T_n^{3/2}}{T_{n+1} - T_n}$ , as  $Y_0$  and  $Y_N$  are fixed in  $K$ .

Following “optimal transport” theory, we may introduce on  $H$  the quadratic Monge–Kantorovich (MK2) (or “Wasserstein”) distance,

$$d_{MK^2}(X, \tilde{X}) = \inf\{\|X \circ s - \tilde{X} \circ \tilde{s}\|, s, \tilde{s} \in S\}, \tag{4.58}$$

which is nothing but the quotient distance in  $H$  with respect to the action of the semi-group  $S$ . Then the minimization of the time-discrete MAG action is just the minimization in  $Y_n \in K$  of

$$\sum_{n=0}^{N-1} r_n d_{MK^2}(Y_{n+1}, Y_n + (Y_n - \text{Id})\theta_n)^2, \tag{4.59}$$

with  $r_n = \frac{T_n^{3/2}}{T_{n+1} - T_n}$ , as  $Y_0$  and  $Y_N$  are fixed in  $K$ .

(Equivalently, we could work on the so-called “Wasserstein” or “MK2” space as, for instance, in [37, 3, 2].)

### 4.3. The fully discrete least action principle

Let us now introduce a fully discrete scheme, for which not only the time variable but also the space variable is discrete. The domain is divided into  $L$  disjoint subdomains  $D_i$  of Lebesgue measure  $1/L$ , for  $i = 1, \dots, L$ , with barycenter  $a_i$  and vanishing diameter as  $L \rightarrow \infty$ . In our abstract framework, it is enough to substitute for the spatial domain  $D$ , the discrete set  $\{a_i, i = 1, \dots, L\}$ . Accordingly,  $H$  can be seen as the Euclidean space  $(\mathbb{R}^d)^L$  of all finite sequences of  $L$  points in  $\mathbb{R}^d$   $\{X = (X_i \in \mathbb{R}^d)_{i=1,L}\}$  with the natural Euclidean norm  $\|\cdot\|$  induced by  $\mathbb{R}^d$ . Meanwhile the set  $S$  can be viewed as the set of all permutations  $s$  of the  $L$  first integers and  $K$  is the corresponding cone of all sequences  $Y_i$  such that

$$\sum_i Y_i \cdot (a_i - a_{s_i}) \geq 0,$$

for all permutations  $s$ . In one space dimension,  $K$  is just the convex cone of all increasing sequences of  $L$  real numbers.

The time-discrete MAG action (4.57) makes sense at the fully discrete level without modification. In this discrete setting,  $S$  is a group (which is untrue at the continuous level) and each  $s$  can be inverted in  $S$ . Thanks to the group property of  $S$  and the invariance of  $\|\cdot\|$  with respect to  $S$ , the minimization problem can be further reduced to the minimization of

$$\sum_{n=0}^{N-1} r_n \|Y_{n+1} - (Y_n + (Y_n - \text{Id})\theta_n) \circ \sigma_{n+1}\|^2, \tag{4.60}$$

in  $Y_n \in K, \sigma_n \in S$ , as  $Y_0$  and  $Y_N$  are fixed in  $K$  (just by setting  $\sigma_{n+1} = s_n \circ s_{n+1}^{-1} \in S$ ).

To solve this minimization problem, a crude strategy is to use Gauss–Seidel type iterations. We denote by  $(Y_n^k, \sigma_n^k)$  the approximation of  $(Y_n, \sigma_n)$  at iteration  $k$  and time step  $n$ . Let us fix  $k$  and  $n$ . To get the updated values  $\sigma_n^{k+1}$  and  $Y_n^{k+1}$ , we inductively suppose that we already know  $(Y_m^j, \sigma_m^j)$  for all  $m$  if  $j \leq k$  and for all  $m < n$  if  $j = k + 1$ . Then, we perform the following two steps:

- (i) First step: we get  $s = \sigma_n^{k+1}$  by solving the combinatorial optimization problem

$$\inf_{s \in S} \|Y_{n+1}^k - (Y_n^k + (Y_n^k - \text{Id})\theta_n) \circ s\|. \tag{4.61}$$

This step is particularly simple in one space dimension and just amounts to sorting in increasing order the finite sequence  $(Y_n^k + (Y_n^k - \text{Id})\theta_n)_i, i = 1, \dots, L$ . It is much more challenging in higher dimensions. The best known optimization methods need  $O(L^3)$  elementary operations, which is not satisfactory (see a related discussion in [18]).

- (ii) Second step: we get  $Y = Y_n^{k+1}$  by minimizing in  $Y \in K$ :

$$\begin{aligned} r_n \|Y_{n+1}^k - (Y + (Y - \text{Id})\theta_n) \circ \sigma_{n+1}^k\|^2 + r_{n-1} \|Y \\ - (Y_{n-1}^{k+1}(1 + \theta_{n-1}) - \theta_{n-1}\text{Id}) \circ \sigma_n^{k+1}\|^2, \end{aligned}$$

where the first term can also be written

$$r_n \|Y_{n+1}^k \circ (\sigma_{n+1}^k)^{-1} - (Y + (Y - \text{Id})\theta_n)\|^2,$$

using the inverse permutation  $(\sigma_{n+1}^k)^{-1}$  and the invariance of  $\|\cdot\|$  with respect to permutations. After reorganizing squares, we see that  $Y$  is just the least-square projection  $H \rightarrow K$  of:

$$V = \frac{r_n(1 + \theta_n)W + r_{n-1}Z}{r_n(1 + \theta_n)^2 + r_{n-1}},$$

$$W = Y_{n+1}^k \circ (\sigma_{n+1}^k)^{-1} + \theta_n \text{Id}, \quad Z = (Y_{n-1}^{k+1} + (Y_{n-1}^{k+1} - \text{Id})\theta_{n-1}) \circ \sigma_n^{k+1}.$$

So, we have obtained an effective algorithm. It is particularly simple in one space dimension (and much more challenging in higher dimensions!). Let us observe that, in one space dimension, computing the least-square projection  $Y = P_K[V]$  is different from sorting the sequence  $V$  in increasing order. However, still in one space dimension, this projection can be approximately computed after a sequence of sorting steps, according to the asymptotic formula (which is a by-product of the “transport-collapse method” [11]):

$$P_K[V] = \lim_{M \rightarrow \infty} V_M^M, \quad V_m^M = \left( V_{m-1}^M + \frac{1}{M} V \right)^\sharp, \quad V_0^M = 0, \quad m = 1, \dots, M. \tag{4.62}$$

In practice, we already get a good accuracy for moderate values of  $M$  (say  $M = 10$ ).



### 5. Solution of the Initial Value Problem

In order to validate the reconstruction scheme, we would like to solve the initial value problem (IVP) consistently with the modified least action problem, and get a discrete scheme for the IVP. Ideally, such a scheme should be derived directly from the modified discrete least action principle. Unfortunately, we have not been able to do so, and we are just going to suggest a simple scheme for the IVP which seems, in practice, consistent with the modified action, at least in one space dimension.

#### 5.1. A time-discrete scheme for the IVP

Our suggestion to get a time-discrete solution of the IVP is to alternate the solution of the linear ODE

$$\frac{d}{dt}(\alpha^2(t)V) = t^{-2}\beta^2(t)(Y - \text{Id}), \quad \frac{dY}{dt} = V, \quad Y(T_n) = Y_n, \quad V(T_n) = V_n, \tag{5.63}$$

with coefficients  $(\alpha, \beta)$  given by (0.4), on each time interval  $[T_n, T_{n+1}[$  and the rearrangement of the result at time step  $T_{n+1}$ :

$$Y_{n+1} = Y(T_{n+1})^\sharp, \quad V_{n+1} = V(T_{n+1}).$$

Using a plain explicit discretization of (5.63), we get:

$$\begin{aligned} Y_{n+1} &= (Y_n + (T_{n+1} - T_n)V_n)^\sharp, \\ \alpha^2(T_{n+1})V_{n+1} &= \alpha^2(T_n)V_n + T_n^{-2}\beta^2(T_n)(T_{n+1} - T_n)(Y_n - \text{Id}). \end{aligned} \tag{5.64}$$

The convergence analysis of this time-discrete scheme can be done in two different ways.

#### 5.2. The multidimensional case

In the multidimensional case, our strategy for the convergence analysis of scheme (5.64) is inspired by our recent work [16], where a similar scheme is analyzed. We essentially use the fact that all maps with convex potential are of locally bounded variations, which provides enough compactness with respect to space variables. Time compactness is, as usual, directly obtained from the evolution scheme. We notice that  $V$  can be easily integrated out from  $Y$  by ODE (5.63).

**Theorem 5.1.** *For every fixed initial condition  $(Y_0, V_0) \in K \times H$ , the approximate solution  $(Y_n)$  admits at least a limit  $t \rightarrow Y(t) \in K$  valued in  $C^0([t_0, +\infty[, H)$  as the time step goes to zero. This limit satisfies:*

$$\begin{aligned} V(t, a)\alpha^2(t) &= V_0(a)\alpha^2(t_0) + \int_{t_0}^t \tau^{-2}\beta^2(\tau)(Y(\tau, a) - a)d\tau, \\ \frac{d}{dt} \int_D f(Y(t, a))da &= \int_D (\nabla f)(Y(t, a)) \cdot V(t, a)da, \end{aligned} \tag{5.65}$$

for all  $C^1$  function  $f$  on  $\mathbb{R}^d$  (with  $|\nabla f|$  growing at most linearly at infinity).

Notice that, since  $Y$  is valued in  $K$ , the knowledge of “observables”  $\int_D f(Y(t, a))da$  for all test-functions  $f$  is enough to determine  $Y(t)$  which makes formulation (5.65) self-consistent. (However, this does not guarantee uniqueness of solutions to the IVP.) So, we have a proposal to solve the IVP, and a corresponding discrete scheme, but we are not able to prove that formulation (5.65) is actually consistent with our modified least action principle.

### 5.3. The one-dimensional case

In the special case of one space variable, we get a much more precise information, following the analysis developed in [14] for similar problems (see also [15]):

**Theorem 5.2.** *For every fixed initial condition  $(Y_0, V_0) \in K \times H$ , as the time step goes to zero, the approximate solution  $(Y_n)$  converges to the unique solution  $t \rightarrow Y(t) \in K$ , valued in  $C^0([t_0, +\infty, H])$ , of the mixed integral-sub-differential system:*

$$\begin{aligned}
 &-\partial_t Y + V \in \partial\Theta[Y], \\
 &V(t, a)\alpha^2(t) = V_0(a)\alpha^2(t_0) + \int_{t_0}^t \tau^{-2}\beta^2(\tau)(Y(\tau, a) - a)d\tau,
 \end{aligned}
 \tag{5.66}$$

where  $\Theta[Y] = 0$  whenever  $Y = Y(t, a)$  is monotonically increasing in  $a$  and  $\Theta[Y] = +\infty$  otherwise.

System (5.66) is well-posed in the  $L^2$  sense and can be shown (as in [14]) to be the limit (in the sense of maximal monotone operator theory) as  $\epsilon \downarrow 0$  of the perturbed system

$$-\partial_t Y + V = -\epsilon\partial_a(\log(\partial_a Y)), \quad \partial_t(\alpha^2(t)V) = t^{-2}\beta^2(t)(Y - \text{Id}),
 \tag{5.67}$$

which, in Eulerian variables (2.29), reduces to:

$$\begin{aligned}
 &\partial_t \rho + \partial_x(\rho v) = 0, \\
 &\partial_t(\alpha^2(t)\rho v) + \partial_x(\alpha^2(t)\rho v^2) = -t^{-1}\beta^2(t)\rho\partial_x \varphi + \epsilon\partial_{xx} v, \\
 &\rho = 1 + t\partial_x^2 \varphi,
 \end{aligned}
 \tag{5.68}$$

and is just a pressure-less Navier–Stokes–Poisson system with vanishing viscosity (as in [10, 42]). As mentioned above, an interesting open question is to show that this approach (vanishing viscosity (5.68), subdifferential formulation (5.66) or scheme (5.64)) is actually consistent with the modified least action principle!

## 6. Numerical Simulations in One Space Dimension

Our data are

$$\begin{aligned}
 t_0 = 1/2, \quad t_1 = 5/2, \quad N = 60, \quad L = 51, \quad a_i = -1 + (2i - 1)/L, \quad i = 1, \dots, L, \\
 Y_0 = X_0^\sharp, \quad (X_0)_i = a_i\omega_i,
 \end{aligned}$$

where  $\omega_i$  is a random number uniformly distributed between 1 and 2. Thus,  $Y_0$  looks like a devil’s staircase.

Concerning the final data  $Y_N \in K$ , either:

(Case 1) the associate probability  $\rho_N$  is the barycenter of four Dirac's measures:

$$\rho_N(dx) = \frac{\delta(x + 0.7) + 4 \delta(x - 0.2) + 3 \delta(x - 0.9) + \delta(x - 1.1)}{9}$$

or (Case 2)  $Y_N$  is the solution at time  $t_1$  of the initial value problem generated by the discrete gradient flow equation starting from  $Y_0 = X_0^\sharp$  at time  $t_0$ .

In our plots, we draw the trajectories of the 51 particles during the 60 time steps of the time interval (the vertical axis corresponding to time and the horizontal one to space).

**Case 1.** We first plot the reconstructed solution (Fig. 1). Then, with the reconstructed initial velocity, we solve the initial value problem for the MAG equations with scheme (5.64) and plot the result (Fig. 2). We observe a nearly perfect match between Figs. 1 and 2.

**Case 2.** We first solve the IVP for the gradient flow equation with scheme (5.64) (Fig. 3). Then, we reconstruct the solution from the initial and final data of the gradient flow solution (Fig. 4). (Here we observe some limited discrepancy.) Finally, with the reconstructed initial velocity we solve the initial value problem for the MAG equations (Fig. 5), again with scheme (5.64) and get a nearly perfect match.

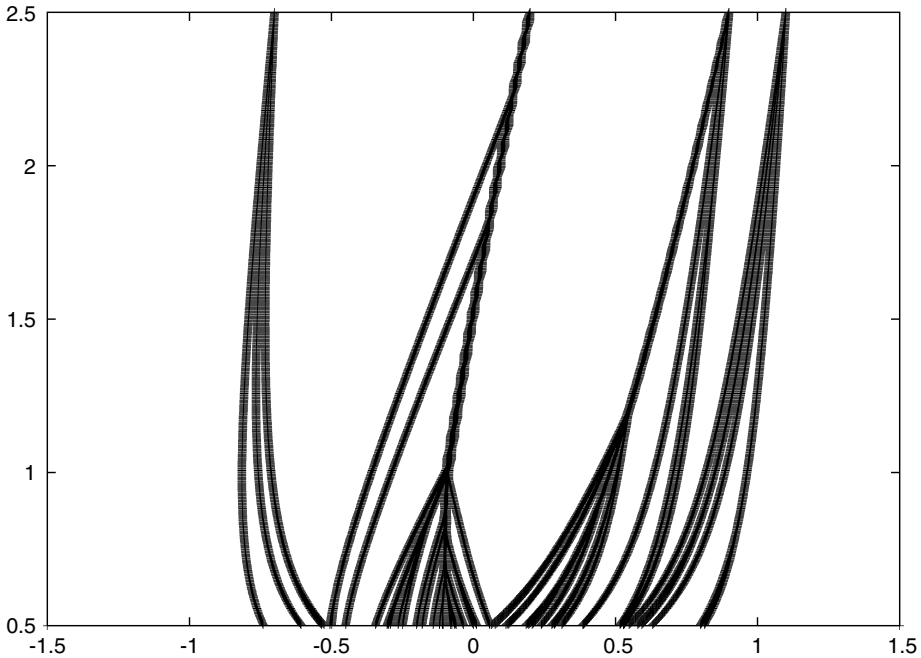


Fig. 1. Case 1/reconstruction.

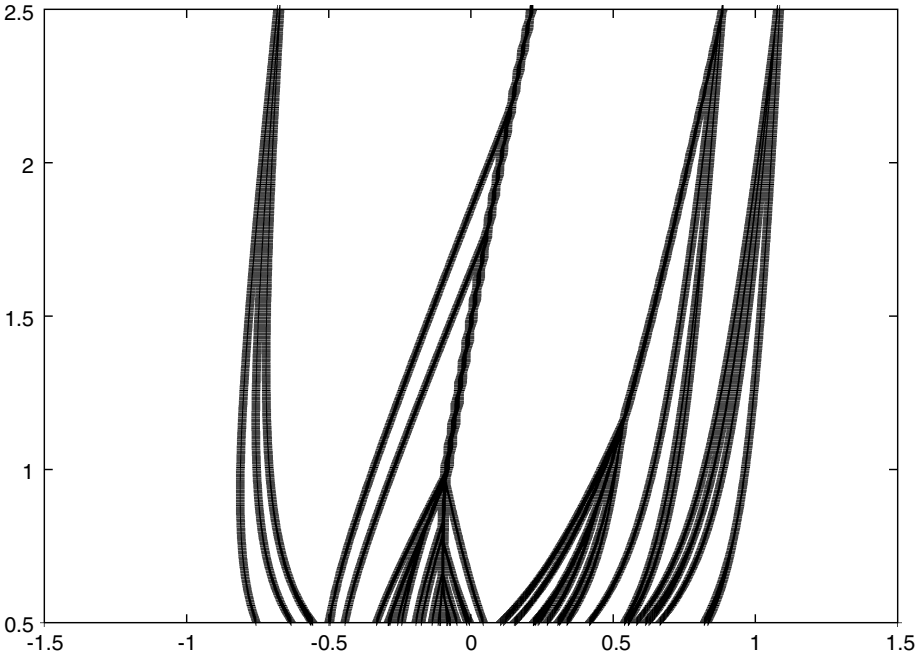


Fig. 2. Case 1/initial value problem (IVP) after reconstruction.

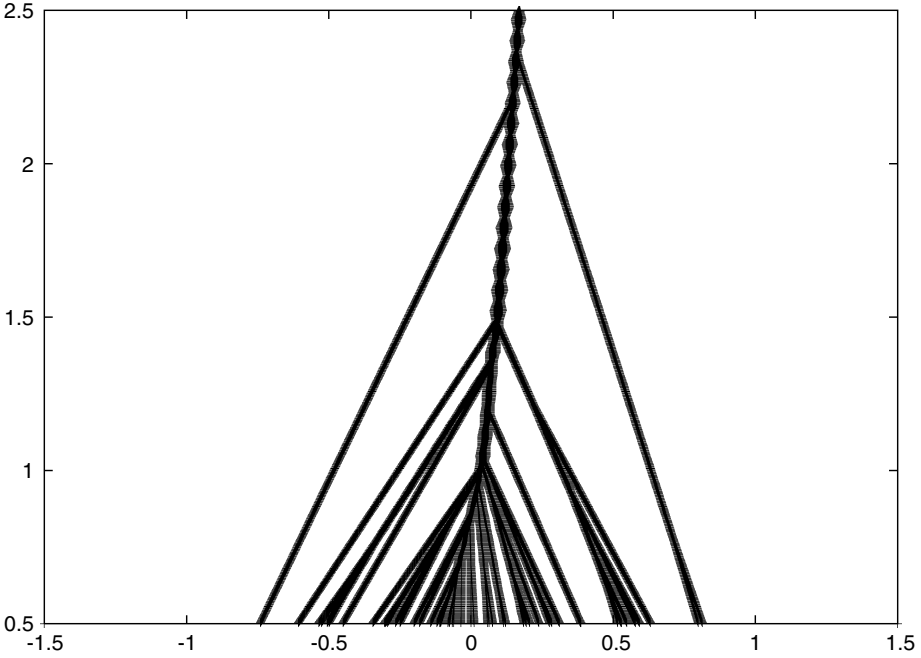


Fig. 3. Case 2/gradient flow solution, IVP before reconstruction.

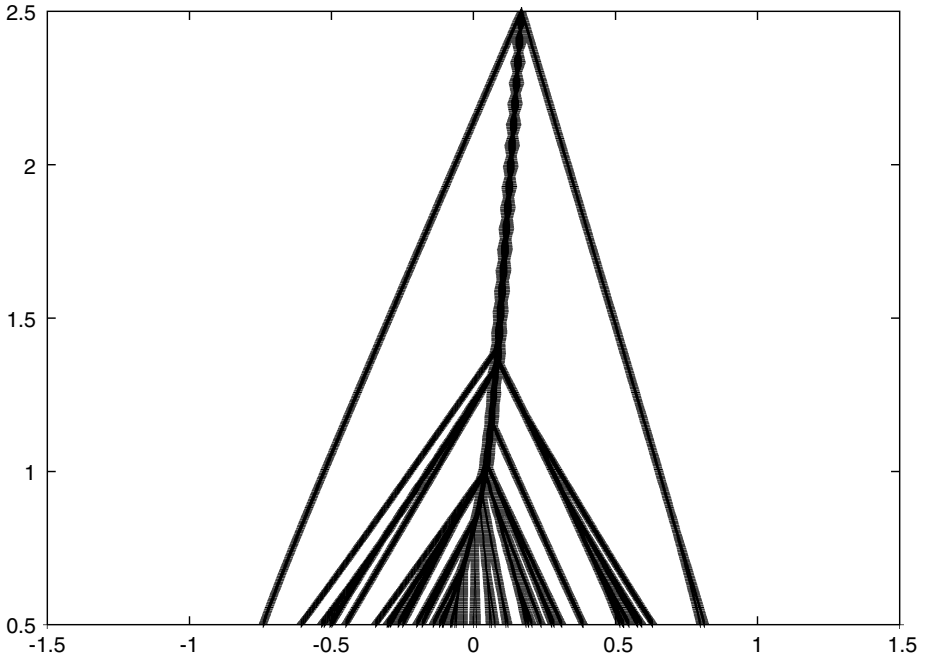


Fig. 4. Case 2/gradient flow solution, reconstruction.

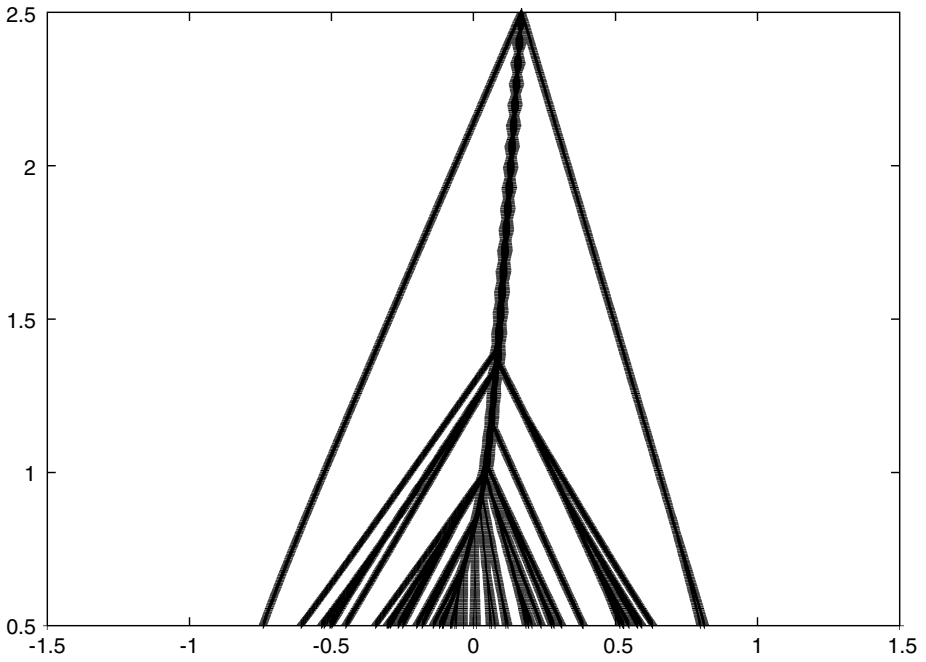


Fig. 5. Case 2/gradient flow solution, IVP after reconstruction.

### 7. Discussion

We have revisited the early universe reconstruction problem and suggested a modification of classical Newton gravitation by what we called Monge–Ampère gravitation. The main drawback of our approach is the lack of physical justification for such a modification. The main mathematical advantage is the obtention of a modified least action principle in which we can easily include mass concentration effects in an almost canonical way, using ideas from gradient flow theory. In addition, the well-known Zeldovich approximate solutions turn out to be exact solutions of the modified model, which provides an indirect validation of the model as a reasonable approximation for the early universe reconstruction (EUR) problem. According to these ideas, an algorithm has been designed in the 1D case. Our plan for the future includes: (i) analysis of the initial value problem, consistently with the modified least action principle; (ii) design of an efficient multidimensional algorithm; (iii) study of the relative accuracy of the Newton and Monge–Ampère gravitation models with respect to general relativity.

### Acknowledgments

The author acknowledges the support of ANR contract OTARIE ANR-07-BLAN-0235. He warmly thanks the hospitality of the Forschungsinstitut Mathematik of the ETHZ that he visited in the fall of 2009 and the Isaac Newton Institute in the summer of 2010.

### Appendix

#### A.1. Some useful results from optimal transport theory

The set  $S$  defined by (2.24) has a semigroup structure for the composition rule and has the identity map  $\text{Id}$  as neutral element. It is, in some sense, in duality with its “polar cone”  $K \subset H$ :

$$K = \{Y \in H; ((Y, \text{Id} - s)) \geq 0, \forall s \in S\}. \tag{A.69}$$

Let us recall a few basic results of optimal transport theory [12, 13, 44] concerning  $S$  and  $K$ . First, the set  $K$  can be characterized as the closed convex cone of all maps  $Y$  with a convex potential, which means that there is a convex function  $\psi$  defined on  $\mathbb{R}^d$  and valued in  $]-\infty, +\infty]$  which is almost everywhere differentiable on  $D$  with  $\nabla\psi(x) = Y(x)$ , a.e. on  $D$ .

Next, every map admits a unique rearrangement in  $K$ . More precisely:

**Theorem A.1.** ([12]) *Every  $X \in H$  admits a unique “rearrangement”  $X^\#$  in  $K$ , which means:*

$$\int_D f(X^\#(a))da = \int_D f(X(a))da, \quad \forall f \in C(\mathbb{R}^d), \quad \sup_x |f(x)|(1 + |x|^2)^{-1} < +\infty.$$

*In addition,  $X \rightarrow X^\#$  is continuous in  $H$  (for the strong topology).*

Moreover, there is a “polar factorization” of the Hilbert space  $H$  by  $S$  and  $K$ . More precisely:

**Theorem A.2.** ([12]) *Let  $X \in H$  be a nondegenerate map, in the sense that the measure  $\rho(dx) = \int_D \delta(x - X(a))da$  has no singular part with respect to the Lebesgue measure.*

*Then  $X$  admits a unique “polar factorization”*

$$X = Y \circ s, \quad Y \in K, \quad s \in S. \tag{A.70}$$

*In addition, the second factor  $s$  is characterized as the unique closest point  $\pi[X]$  to  $X$  in  $S$  and can be written*

$$\pi[X] = \nabla\psi \circ X, \tag{A.71}$$

*where  $\nabla\psi$  is the unique map  $T: \mathbb{R}^d \rightarrow D$  with convex Lipschitz potential such that the Lebesgue measure restricted to  $D$  is the image of  $\rho$  by  $T$ :*

$$\int_{\mathbb{R}^d} f(\nabla\psi(x))\rho(dx) = \int_D f(a)da, \quad \forall f \in C(\mathbb{R}^d). \tag{A.72}$$

Let us finally observe as in [12, 13] that (A.72) can be seen as a “weak formulation” (not in the sense of distributions!) of the Monge–Ampère problem on  $\mathbb{R}^d$  with range condition:

$$\rho = \det(D_x^2\psi), \quad (\nabla\psi)(\mathbb{R}^d) = D. \tag{A.73}$$

## References

1. L. Ambrosio, Transport equation and Cauchy problem for  $BV$  vector fields, *Invent. Math.* **158** (2004) 227–260.
2. L. Ambrosio and W. Gangbo, Hamiltonian ODE in the Wasserstein spaces of probability measures, *Comm. Pure Appl. Math.* **61** (2008) 18–53.
3. L. Ambrosio, N. Gigli and G. Savaré, *Gradient Flows in Metric Spaces and the Wasserstein Spaces of Probability Measures*, Lectures in Mathematics (Birkhäuser, 2005).
4. J.-P. Aubin, *Mathematical Methods of Game and Economic Theory*, Studies in Mathematics and Its Applications, Vol. 7 (North-Holland, 1979).
5. M. Avellaneda and W. E. Statistical properties of shocks in Burgers turbulence, *Comm. Math. Phys.* **172** (1995) 13–38.
6. E. Aurell, U. Frisch, J. Lutsko and M. Vergassola, On the multifractal properties of the energy dissipation derived from turbulence data, *J. Fluid Mech.* **238** (1992) 467–486.
7. M. Blaser and T. Rivière, A minimality property for entropic solutions to scalar conservation laws in  $1 + 1$  dimensions, arXiv:0907.4215v1.
8. F. Bouchut, *Advances in Kinetic Theory and Computing*, Ser. Adv. Math. Appl. Sci., Vol. 22 (World Scientific, 1994).
9. F. Bouchut, Renormalized solutions to the Vlasov equation with coefficients of bounded variation, *Arch. Rational Mech. Anal.* **157** (2001) 75–90.
10. L. Boudin, A solution with bounded expansion rate to the model of viscous pressureless gases, *SIAM J. Math. Anal.* **32** (2000) 172–193.

11. Y. Brenier, Averaged multivalued solutions for scalar conservation laws, *SIAM J. Numer. Anal.* **21** (1984) 1013–1037.
12. Y. Brenier, Décomposition polaire et réarrangement monotone des champs de vecteurs, *C. R. Acad. Sci. Paris I Math.* **305** (1987) 805–808.
13. Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions, *Comm. Pure Appl. Math.* **44** (1991) 375–417.
14. Y. Brenier, Order preserving vibrating strings and applications to electrodynamics and magnetohydrodynamics, *Methods Appl. Anal.* **11** (2004) 515–532.
15. Y. Brenier,  $L_2$  formulation of multidimensional scalar conservation laws, *Arch. Rational Mech. Anal.* **193** (2009) 1–19.
16. Y. Brenier, Optimal transport, convection, magnetic relaxation and generalized Boussinesq equations, *J. Nonlinear Sci.* **19** (2009) 547–570.
17. Y. Brenier and E. Grenier, Sticky particles and scalar conservation laws, *SIAM J. Numer. Anal.* **35** (1998) 2317–2328.
18. Y. Brenier, U. Frisch, M. Hénon, G. Loeper, S. Matarrese, R. Mohayaee and A. Sobolevskii, Reconstruction of the early universe as a convex optimization problem, *Mon. Not. R. Astron. Soc.* 2002.
19. Y. Brenier and G. Loeper, A geometric approximation to the Euler equations: The Vlasov–Monge–Ampère equation, *Geom. Funct. Anal.* **14** (2004) 1182–1218.
20. H. Brezis, *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert*, North-Holland Mathematics Studies, No. 5 (North-Holland, 1973).
21. J. Carrillo, M. Di Francesco, A. Figalli, T. Laurent and D. Slepcev, Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations, to appear in *Duke Math. J.*, CVGMT preprint 2010.
22. D. Christodoulou, *The Action Principle and PDEs*, Annals of Maths Studies, Vol. 146 (Princeton Univ. Press, 2000).
23. M. Cullen and J. Purser, An extended Lagrangian theory of semigeostrophic frontogenesis, *J. Atmospheric Sci.* **41** (1984) 1477–1497.
24. C. Dafermos, *Hyperbolic Conservation Laws in Continuum Physics* (Springer, Berlin, 2000).
25. W. E, Y. Rykov and Y. Sinai, Generalized variational principles, global weak solutions and behavior with random initial data for systems of conservation laws arising in adhesion particle dynamics, *Comm. Math. Phys.* **177** (1996) 349–380.
26. U. Frisch, S. Matarrese, R. Mohayaee and A. Sobolevski, A reconstruction of the initial conditions of the universe by optimal mass transportation, *Nature* **417** (2002) 260–262.
27. W. Gangbo, T. Nguyen and A. Tudorascu, Hamilton–Jacobi equations in the Wasserstein space, *Methods Appl. Anal.* **15** (2008) 155–183.
28. W. Gangbo, T. Nguyen and A. Tudorascu, Euler–Poisson systems as action-minimizing paths in the Wasserstein space, *Arch. Rational Mech. Anal.* **192** (2009) 419–452.
29. N. Ghoussoub, *Self-Dual Partial Differential Systems and Their Variational Principles* (Springer, 2009).
30. G. Gibbons and S. Hawking, *The Very Early Universe* (Cambridge Univ. Press, 1983).
31. W. Jäger and S. Luckhaus, On explosions of solutions to a system of partial differential equations modelling chemotaxis, *Trans. Amer. Math. Soc.* **329** (1992) 819–824.
32. E. Keller and L. Segel, Model for chemotaxis, *J. Theor. Biol.* **30** (1971) 225–234.
33. G. Loeper, The reconstruction problem for the Euler–Poisson system in cosmology, *Arch. Rational Mech. Anal.* **179** (2006) 153–216.



34. L. Natile and G. Savaré, A Wasserstein approach to the one-dimensional sticky particle system, *SIAM J. Math. Anal.* **41** (2009) 1340–1365.
35. T. Nguyen and A. Tudorascu, Pressureless Euler/Euler–Poisson systems via adhesion dynamics and scalar conservation laws, *SIAM J. Math. Anal.* **40** (2008) 754–775.
36. J. Nieto, F. Poupaud and J. Soler, High-field limit for the Vlasov–Poisson–Fokker–Planck system, *Arch. Rational Mech. Anal.* **158** (2001) 29–59.
37. F. Otto, The geometry of dissipative evolution equations, *CPDE* **26** (2001) 873–915.
38. P. J. E. Peebles, *Astrophys. J.* **344** (1989) L53–L56.
39. M. Sever, An existence theorem in the large for zero-pressure gas dynamics, *Diff. Integral Eq.* **14** (2001) 1077–1092.
40. S. Shandarin and Y. Zeldovich, The large-scale structure of the universe: Turbulence, intermittency, structures in a self-gravitating medium, *Rev. Mod. Phys.* **61** (1989) 185–220.
41. A. Shnirelman, On the principle of the shortest way in the dynamics of systems with constraints, in *Global Analysis Studies and Applications*, Vol. II, Lecture Notes in Math., Vol. 1214 (Springer, 1986), pp. 117–130.
42. A. Sobolevskii, The small viscosity method for a one-dimensional system of equations of gas dynamic type without pressure, *Dokl. Akad. Nauk* **356** (1997) 310–312.
43. M. Vergassola, B. Dubrulle, U. Frisch and A. Noullez, Burgers’ equation, Devil’s staircases and the mass distribution for large-scale structures, *Astron. Astrophys.* **289** (1994) 325–356.
44. C. Villani, *Topics in Optimal Transportation*, Graduate Studies in Mathematics, Vol. 58 (Amer. Math. Soc., 2003).
45. G. Wolansky, On time reversible description of the process of coagulation and fragmentation, *Arch. Rational Mech. Anal.* **193** (2009) 57–115.
46. Y. Zeldovich, Gravitational instability: An approximate theory for large density perturbations, *Astron. Astrophys.* **5** (1970) 84–89.