Structured population models of cell differentiation - stability in the space of Radon measures

Grzegorz Jamróz

Preprint no. 2012 - 031

INNOVATIVE ECONOMY
NATIONAL COHESION STRATEGY

FNP
Foundation for Polish Science

EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND

Ph.D. Programme: Mathematical Methods in Natural Sciences (MMNS)
e-mail: mmns@mimuw.edu.pl
http://mmns.mimuw.edu.pl

Structured population models of cell differentiation – stability in the space of Radon measures

Grzegorz Jamróz

Institute of Applied Mathematics and Mechanics,
University of Warsaw, Banacha 2, 02-097 Warszawa, Poland
e-mail: jamroz@mimuw.edu.pl

October 18, 2012

Abstract

We investigate the notion of stability of structured population models of cell differentiation. Metrics in the space of Radon measures on $\mathbb{R}$ prove to be a convenient tool. We discuss their general structure and select a metric which is appropriate for a framework of models based on nonlinear transport equation with measure-valued solutions and Measure-Transmission conditions. We prove the stability of solutions with respect to perturbation of the initial condition. We discuss analytical, technical and modelling consequences of the chosen metric.

Keywords: population dynamics, nonlinear transport equation, cell differentiation, measure-valued solution, transmission conditions, measure-transmission metric

MOS Subject Classification: 28A33; 35F16; 35F31; 92D25

1 Introduction

Mathematical models of cell differentiation[1] include various classes, e.g. coupled systems of ODEs (see e.g. [14]), transport equation (see e.g. [11]), delay differential equations (see e.g. [11]), stochastic models (see e.g. [17]) etc. Here we focus on the general framework of models of transport type introduced in [6]. The framework, which can model various processes, in particular neurogenesis (compare [10]), consists of the following system of equations ([6])

$$\partial_t \mu(t) + \partial_x (g_1(v(t)) \mathbf{1}_{x \neq x_i}(x) \mu(t)) = p(v(t), x)\mu(t), \tag{1}$$

$$g_1(v(t)) \frac{D\mu(t)}{D\mathcal{L}^1}(x_i^+) = c_i(v(t)) \int_{\{x_i\}} d\mu(t), \qquad i = 0, \dots, N \tag{2}$$

$$\mu(0) = \mu_0, \tag{3}$$

where $\mu(t)$ is a measure-valued function, $\mu : \mathbb{R}^+ \to \mathcal{M}(\mathbb{R})$ (Radon measures on $\mathbb{R}$, see [5]), $x_0 < x_1 < \cdots < x_N$ are given points in $\mathbb{R}$ and $\frac{D\mu}{D\mathcal{L}^1}$ denotes the density of the measure μ with respect to one-dimensional

[1] See e.g. [9] for biological introduction to cell differentiation processes

Lebesgue measure. $v(t) = \int_{\{x_N\}} d\mu(t)$ denotes the mass of the last point and is responsible for the feedback(biologically)/nonlinearity(mathematically). $\mathbf{1}_{x \neq x_i}$ is equal 0 if $x \in \{x_0, \ldots, x_N\}$ and 1 otherwise. The initial data are defined on the interval $[x_0, x_N]$, what means that $\mu(0)(A) = 0$ if $A \cap [x_0, x_N] = \emptyset$. g_1, c_i are given nonnegative Lipschitz continuous functions, $c_N \equiv 0$ and $p(v, x) = p_1(v)p_2(x)$ is a product of a Lipschitz continuous function p_1 and a bounded function p_2, which is Lipschitz continuous on open intervals (x_i, x_{i+1}) (compare Assumptions 3.2 in [6]).

This system of equations models the process of cell differentiation, in which both discrete and continuous states, described by structure parameter x, are possible. The values of this parameter correspond to *differentiation level* of a cell. In the model, a cell starting from a given value x may only become more mature, i.e. increase its differentiation level. Discrete/quasistationary states correspond to states where a cell can halt its differentiation for a given period of time. This is not possible in continuous/transient states. Mathematically, cell differentiation consists in traversing the stages x in one direction (to the right) with a velocity $g = g_1(v)\mathbf{1}_{x \neq x_i}$ (equation (1)), which vanishes in discrete (quasistationary) states $x_0, \ldots, x_N$. The measure transmission conditions, given by (2), describe the behavior of the population in those states. They express quantitatively the probability of differentiation from a quasistationary state and may be regarded as constitutive relations of a given cellular system. All $x \in \mathbb{R} \backslash \{x_0, \ldots, x_N\}$ are continuous (transient) states. The biological as well as mathematical motivations are discussed in detail in [6] where also the existence and uniqueness of solutions was demonstrated. The question of stability is addressed in the present article.

The paper consists of three parts:

1. In Section 2 we motivate and derive the proper notion of stability. We discuss the general structure of metrics on the space of Radon measures and select a metric which reflects the desired notion.

2. In Section 3 we prove the stability of the system with respect to perturbation of initial condition in conservative case ($p = 0$). We discuss the analytical and technical consequences of the structure of the metric.

3. In Section 4 we discuss the result and describe the possible applications.

1.1 Notation and terminology

Below, $j \in \{1, 2\}$ and enumerates two solutions $\mu_j(t)$.

- $v_j(t) := \int_{\{x_N\}} d\mu_j(t)$
- $G_j(t) := \int_0^t g_1(v_j(s))ds$
- $\min(G_j) := \min(G_1(t), G_2(t))$
- $\max(G_j) := \max(G_1(t), G_2(t))$
- $J_{min} = J_{min}(T) := [x_N - \min(G_j(T)), x_N)$
- $J_{max} = J_{max}(T) := (x_N - \max(G_j(T)), x_N)$
- $\tau_j(x)$ is the unique solution of the equation $x = x_i - \int_0^{\tau_j(x)} g_1(v_j(s))ds$ for some i, whose value depends on the quasistationary point dealt with.

- $\tau_{\min}(x) := \min(\tau_1(x), \tau_2(x))$
- x_b denotes the argument of measures $\mu_j(0)$
- $\min(g_1)$ denotes the minimum of $g_1(v(t))$ on the interval considered, i.e. on $[0, T]$ (if not stated explicitly otherwise).
- $\text{Lip}(g_1), \text{Lip}(c_i)$ etc. denote the respective global Lipschitz constants; $\sup(c_i), \sup(\psi)$ etc. mean in fact global suprema of the absolute value, i.e. $\sup(|c_i|), \sup(|\psi|)$. Other similar definitions are self-explanatory. The only exception is $\min(g_1)$ (see above).
- $\text{Lip}(c) = \max_{i\in\{0\dots N\}} \text{Lip}(c_i)$; $\sup(c) = \sup_{i\in\{0\dots N\}} \sup(c_i)$
- $\text{Lip}(\psi) = \max_{i\in\{0\dots N\}} \text{Lip}(\psi\lfloor_{(x_{i-1},x_i]})$
- ψ is a test function from $B^{1,\infty}_{MT}$, which is used at time T; ψ^0 denotes the pullback of ψ to time 0 along the flow generated by $g_1(v_1)$ and $c_i(v_1)$.
- $\lceil x \rceil$ stands for the smallest integer greater or equal x.
- By characteristic starting in x_0 we mean any solution $x_j(t)$ of equation
$$x_j(t) - x_0 = \int_0^t g_1(v_j(s))\mathbf{1}_{x_j(s)\neq x_i} ds.$$
- $T_{int} := \frac{|x_N - x_{N-1}|}{\sup(g_1)}$
- $T_{max} := \min_{i\in\{1,\dots,N-1\}} \frac{|x_i - x_{i+1}|}{\sup(g_1)}$
- $\mathbb{R}$ denotes the set of real numbers.
- $\text{Lip}(A)$ is the Banach space of Lipschitz-continuous functions on set A with norm $\|f\| = \max(\sup(f), \text{Lip}(f))$, where $\text{Lip}(f)$ denotes the Lipschitz constant of function f.
- $\mathcal{B}(\mathbb{R})$ stands for the set of Borel-measurable functions on $\mathbb{R}$.
- $BV_{loc}([0,\infty))$ is the set of locally BV functions (functions with bounded variation) on $[0,\infty)$ (see [5]).
- $\mathcal{M}(\mathbb{R})$ stands for the set of all Borel, locally finite measures on $\mathbb{R}$, i.e. Radon measures (see [5]).
- $C^0(\mathbb{R})$ is the set of all bounded continuous functions on $\mathbb{R}$.
- $C([0,\infty), metricspace)$ is a space of all functions defined on $[0,\infty)$ and with values in the given metric space, which are continuous.
- $supp(\mu)$ denotes the support of measure μ.
- x_{-1} and x_{N+1} are by convention equal $-\infty$ and $+\infty$, respectively.
- $\mathbf{1}_A(x)$ is the characteristic function of set A which is equal 1 if $x \in A$ and 0 if $x \notin A$. In particular $\mathbf{1}_{x\neq x_i} = \mathbf{1}_A$ for $A = \mathbb{R}\backslash\{x_0, x_1, \dots, x_N\}$.
- $Var(\mu) := \int_{\mathbb{R}} d\mu$ is the total variation of measure μ.

1.2 Auxiliary estimates

Estimate 1. $|e^x - 1| \le |x|e^x$.

Estimate 2. $|e^{-x} - e^{-y}| \le |x-y|e^{-\min(x,y)}$ for $x, y \ge 0$.

Estimate 3.

$$\sup_{t\in[0,T]} \left|e^{\xi(t)} - 1\right| = \sup_{t\in[0,T]} \left|\int_0^{\xi(t)} e^s ds\right| \le \sup_{t\in[0,T]} |\xi(t)|\, e^{\sup_{t\in[0,T]}|\xi(t)|}.$$

Estimate 4. $|\tau_2(x_b) - \tau_1(x_b)| \le \frac{Lip(g_1)}{\min(g_1)} \int_0^T |v_2(s) - v_1(s)|$.

Proof.

$$|\tau_2(x_b) - \tau_1(x_b)| \le \frac{|x_1(\min(\tau_1,\tau_2)) - x_2(\min(\tau_1,\tau_2))|}{\min(g_1)} \le$$

$$\frac{1}{\min(g_1)} \int_0^{\min(\tau_1,\tau_2)} |g_1(v_2(s)) - g_1(v_1(s))| ds \le \frac{Lip(g_1)}{\min(g_1)} \int_0^T |v_2(s) - v_1(s)| ds$$

□

2 Stability of the system

In this section we motivate, discuss and define stability of system (1)–(3) with respect to perturbation of the initial condition. We identify perturbations, which from the experimental point of view are small (and due to measurement errors). Our goal is to prove that the solution of a perturbed system stays in finite time intervals arbitrarily close to the solution of the unperturbed one provided the perturbation is small enough.

2.1 Strong or weak with respect to initial condition?

Metrics metrizing weak[2] convergence in the space of Radon measures are in general well suited for structured population models ([3, 7, 8]). Strong (i.e. norm) topology, on the other hand, has rather undesirable properties (see Remark 2.3). As it will turn out, in the present case, an intermediate topology will prove useful.

First, let us examine the general structure of metrics in spaces of measures. Take μ_1, μ_2 – two finite Borel measures on $\mathbb{R}$. A general class of metrics can be defined as follows.

Definition 2.1.

$$\rho(\mu_1, \mu_2) := \sup_{\psi \in TFS} \int_{\mathbb{R}} \psi d(\mu_1 - \mu_2).$$

TFS (Test Function Space) is a given subspace of Borel measurable, everywhere defined functions on $\mathbb{R}$. The most important examples from our point of view are summarized in table 1.

Additional conditions imposed on the class of test functions account for the fact that points x are related to one another. Usually, these conditions are local (defined by some local smoothness property, e.g. Lipschitz condition) and have a straightforward interpretation of (properly understood) distance of points.

[2]Throughout this article by *weak* convergence of measures we mean in fact *weak** convergence, i.e. we write $\mu^n \rightharpoonup \mu$ if for every $f \in C^0(\mathbb{R})$ we have $\int_{[x_0,x_N]} f d\mu^n \to \int_{[x_0,x_N]} d\mu$.

Table 1: Weak metrics in the space of Radon measures and their test function spaces

Name	**Test Function Space (TFS)**	**Metric**	**Strong points**
Norm (strong) distance	$\{f \in \mathcal{B}(\mathbb{R}) : \sup |f| \leq 1\}$	$\| \cdot \|$	All
Measure-Transmission metric	To be determined	ρ_{MT}	Some
1-Wasserstein distance	$\{f \in \mathrm{Lip}(\mathbb{R}) : \mathrm{Lip}(f) \leq 1\}$	ρ_W	None
Bounded Lipschitz distance or flat metric	$\{f \in \mathrm{Lip}(\mathbb{R}) : \mathrm{Lip}(f) \leq 1, \sup |f| \leq 1\}$	ρ_F	None

Definition 2.2. A point $x \in \mathbb{R}$ is a *strong point* of a metric if TFS admits functions with discontinuities (or arbitrarily large slopes) at this point. If only right or left discontinuities are allowed, then x is called *right* or *left strong point*, respectively. Points which admit bounded local slope of test functions are called *weak points.*

Norm distance excludes relationships between points. In this case, all points are strong, what is expressed by TFS being composed of arbitrary Borel functions. At the other extremity we have flat metric and Wasserstein metric, which are related to the optimal transport problem (compare [2, 16]). For them, 1-Lipschitz continuous test functions reflect the euclidean distance of points. The goal of this section is to determine an intermediate TFS which is well adapted to the process of cell differentiation.

Modelling Remark 2.3. *Empirical measurements are usually well described by weak metrics. For instance, suppose that in a cell differentiation model all cells are in the same transient state x. Since it is hard to measure in which one exactly, δ_x should be close to $\delta_{x+\epsilon}$ for small values of ϵ. The norm distance of these Dirac masses equals 2 independently of ϵ. In flat metric (or any other weak metric), however, their distance tends to 0 as $\epsilon \to 0$ due to the fact that flat metric metrizes weak convergence of measures. Therefore, transient states should be weak points. On the other hand, quasistationary points x_i are biologically and biochemically rather distinct from their neighbors. To determine which type of points the quasistationary states x_i should represent we investigate the following two examples.*

Example 2.4. Let us consider the following simplification of model (1)-(3).

$$\begin{aligned} \partial_t \mu(t) + \partial_x(\mathbf{1}_{x \neq -1,0,1}\mu(t)) &= 0 \\ \frac{D\mu(t)}{D\mathcal{L}^1}(x_i^+) &= 0 \qquad i = 0, 1, 2 \end{aligned}$$

This system corresponds to $N = 3$ and three quasistationary points $x_0 = -1, x_1 = 0, x_2 = 0$. The differentiation rate g_1 equals 1 whereas the proliferation rate p as well as transmission coefficients c_0, c_1, c_2 vanish. Consider for $0 < \epsilon < 1$ the initial condition $\mu^\epsilon(0) = \delta_\epsilon$. The solutions are given by formula

$$\mu^\epsilon(t) = \begin{cases} \delta_{\epsilon+t} & \text{for } 0 \leq t \leq 1 - \epsilon, \\ \delta_1 & \text{for } t \geq 1 - \epsilon. \end{cases}$$

For initial condition $\mu^0(0) = \delta_0$, on the other hand, we obtain that $\mu^0(t) = \delta_0$ for every $0 \leq t < \infty$. We can see that $\mu^\epsilon(0)$ converges weakly to $\mu^0(0)$ as $\epsilon \to 0^+$ while $\mu^\epsilon(t)$ converges weakly to

$$\overline{\mu}(t) = \begin{cases} \delta_t & \text{for } 0 \le t \le 1, \\ \delta_1 & \text{for } t \ge 1. \end{cases}$$

Clearly, $\mu^0 \neq \overline{\mu}$.

We conclude that there is *no* weak stability at x_1.

Modelling Remark 2.5. *In fact, Example 2.4 represents a generic case of destabilization. To see this, let us consider the biological context of cell differentiation. δ_0 may describe a stable population of progenitor cells which do not differentiate. If this stable population of cells is pushed out of its stable attractor by an external stimulus, the cells will start to differentiate all at once. Thus, although the structure parameter values $x = 0$ and $x = \epsilon$ are close to each other in euclidean sense, they are not close in the sense of biological system. This is due to the fact that moving from one to the other requires significant work (e.g. significant change of external conditions causing destabilization) and does not happen spontaneously.*

The next example shows, that weak stability from the left is reasonable.

Example 2.6. Consider the same equation and coefficients as in Example 2.4. For initial condition $\mu^{-\epsilon}(0) = \delta_{-\epsilon}$, $0 < \epsilon < 1$, we obtain

$$\mu^{-\epsilon}(t) = \begin{cases} \delta_{-\epsilon+t} & \text{for } t \le \epsilon, \\ \delta_0 & \text{for } t \ge \epsilon. \end{cases}$$

We can see that for every $t \ge 0$ measure $\mu^{-\epsilon}(t)$ converges weakly to $\mu^0(t)$ as $\epsilon \to 0^+$. This means left weak stability at 0.

Modelling Remark 2.7. *Biologically this means that if a stable population of cells is brought back to a slightly less differentiated (i.e. more upstream) state, it will promptly differentiate back to the original state.*

To summarize, transient points should be weakly stable (weak points of a metric) and the quasistationary points – weakly stable 'to the left' (i.e. right strong points).

Corollary 2.8. *From the modeling and simulation point of view, useful data around quasistationary states are such that have a* sharp *distinction between a node point and its downstream neighborhood. Conversely, the upstream neighborhood may be biochemically very similar to the node point. In terms of biochemical markers [3], a useful set of markers is such that distinguishes between progenitor/stem cells and their downstream progeny. Moreover, markers which do not indicate precisely when a cell reaches a quasistationary state are still useful as long as a cell measured by these markers is destined to reach a quasistationary state in a bounded short time.*

2.2 The Measure-Transmission metric

Based on the above considerations, we define the *Measure-Transmission metric* ρ_{MT} on $\mathcal{M}(\mathbb{R})$.

The test function space, $W^{1,\infty}_{MT}$, is composed of functions which are Lipschitz-continuous on intervals $(x_{i-1}, x_i]$ and left-continuous in x_i, $i = 0, \ldots, N$.

[3]Consult any standard textbook on cell/molecular biology, e.g. [15] for an introduction to biological marking techniques

Definition 2.9.

$$W^{1,\infty}_{MT}(\mathbb{R}) := \Big\{\psi \in \mathcal{B}(\mathbb{R}) : \sup|\psi| \leq \infty, \|\psi\lfloor_{(-\infty,x_0]}\|_{\text{Lip}} \leq \infty, \|\psi\lfloor_{(-x_0,x_1]}\|_{\text{Lip}} \leq \infty, \dots, \|\psi\lfloor_{(x_{N-1},x_N]}\|_{\text{Lip}} \leq \infty, \| \psi\lfloor_{(x_N,+\infty)}\|_{\text{Lip}} \leq \infty \Big\}.$$

This space admits a norm.

Definition 2.10.

$$\|\psi\|_{W^{1,\infty}_{MT}} := \max\left(\sup|\psi|, \|\psi\lfloor_{(-\infty,x_0]}\|_{\text{Lip}}, \|\psi\lfloor_{(-x_0,x_1]}\|_{\text{Lip}}, \dots, \|\psi\lfloor_{(x_{N-1},x_N]}\|_{\text{Lip}}, \|\psi\lfloor_{(x_N,+\infty)}\|_{\text{Lip}}\right).$$

With this norm $W^{1,\infty}_{MT}$ is a Banach space as a direct product of finite number of Banach spaces of Lipschitz continuous functions on $(x_{i-1}, x_i]$ for $i \in \{0, \dots, N+1\}$. The unit ball in $W^{1,\infty}_{MT}$ is defined by

Definition 2.11.

$$B^{1,\infty}_{MT}(\mathbb{R}) := \{\psi \in W^{1,\infty}_{MT} : \|\psi\|_{W^{1,\infty}_{MT}} \leq 1\} = \Big\{\psi \in \mathcal{B}(\mathbb{R}) : \sup|\psi| \leq 1, \|\psi\lfloor_{(-\infty,x_0]}\|_{\text{Lip}} \leq 1, \|\psi\lfloor_{(-x_0,x_1]}\|_{\text{Lip}} \leq 1, \dots, \|\psi\lfloor_{(x_{N-1},x_N]}\|_{Lip} \leq 1, \|\psi\lfloor_{(x_N,+\infty)}\|_{Lip} \leq 1 \Big\}.$$

Finally, we define the Measure-Transmission metric, ρ_{MT}.

Let μ_1, μ_2 be two Radon measures on $\mathbb{R}$.

Definition 2.12 (Measure-Transmission metric)**.**

$$\rho_{MT}(\mu_1, \mu_2) := \sup_{\psi \in B^{1,\infty}_{MT}(\mathbb{R})} \int_{\mathbb{R}} \psi d(\mu_1 - \mu_2).$$

Proposition 2.13. *ρ_{MT} is a metric.*

Proof. Symmetry is obvious. Next, $\rho_{MT}(\mu_1, \mu_2) \geq \rho_F(\mu_1, \mu_2)$ and thus $\rho_{MT}(\mu_1, \mu_2) = 0$ iff $\mu_1 = \mu_2$. Finally, from definition 2.12 it is clear that $\rho_{MT}(\mu_1, \mu_3) \leq \rho_{MT}(\mu_1, \mu_2) + \rho_{MT}(\mu_2, \mu_3)$ for any $\mu_1, \mu_2, \mu_3 \in \mathcal{M}(\mathbb{R})$. □

Example 2.14. $\rho_{MT}(\delta_{x_1}, \delta_{x_1+\epsilon}) = 2$*, whereas* $\rho_F(\delta_{x_1}, \delta_{x_1+\epsilon}) = \epsilon$*.*

Proof. In case of ρ_{MT} the supremum from definition 2.12 is realized e.g. by function $\psi = \mathbf{1}_{(-\infty,x_1]} - \mathbf{1}_{(x_1,\infty)}$. In case of ρ_F we cannot use this ψ as test function, since it is not left-continuous in x_1. Instead, to obtain the largest value of $\int_{\mathbb{R}} \psi d(\delta_1 - \delta_{1+\epsilon}) = \psi(1) - \psi(1+\epsilon)$, we take

$$\psi(x) = \begin{cases} 1 & \text{for } x \leq x_1 \\ x_1 - x & \text{for } x \in (x_1 - 1, x_1 + 1) \\ -1 & \text{for } x \geq x_1 + 1. \end{cases}$$

□

Example 2.15. $\rho_{MT}(\delta_{x_0}, \delta_{x_0-\epsilon}) = \rho_F(\delta_{x_0}, \delta_{x_0-\epsilon}) = \epsilon$*.*

Proof. Similar as in the previous example. Note that function $\psi = \mathbf{1}_{(-\infty,x_1)} - \mathbf{1}_{[x_1,\infty)}$ does not belong to $B^{1,\infty}_{MT}$, since it is not left-continuous in x_1. Therefore it cannot be applied as a test function from Definition 2.12. □

These examples express in formulas the fact that a quasistationary node point is 'close' to its own upstream (left) neighbourhood, yet is sharply separated from its downstream (right) neighbourhood. In other words, a quasistationary point is a right strong point of the Measure-Transmission metric.

2.3 Measure-Transmission solutions

Definition 2.16 (Measure-Transmission solution). [4] *A measure-valued function* $\mu \in C([0,\infty),(\mathcal{M},\rho_{MT}))$ *with* $v(t) = \int_{\{x_N\}} d\mu(t) \in BV_{loc}[0,\infty)$ *is called a* Measure-Transmission *solution of problem* (1)–(3)*, if*

(i) for all $\varphi \in C_c^\infty([0,\infty)\times\mathbb{R})$

$$
\begin{aligned}
-\int_{\mathbb{R}^+}\int_{\mathbb{R}} \partial_t\varphi(t,x)d\mu(t)(x)dt - \int_{\mathbb{R}^+}\int_{\mathbb{R}} g_1(v(t))\mathbf{1}_{x\neq x_i}(x)\partial_x\varphi(t,x)d\mu(t)(x)dt \\
= \int_{\mathbb{R}^+}\int_{\mathbb{R}} p_1(v(t))p_2(x)\varphi(t,x)d\mu(t)(x)dt + \int_{\mathbb{R}}\varphi(0,x)d\mu(0)(x),
\end{aligned} \tag{4}
$$

(ii) For $\mathcal{L}^1$ *a.e.* $t \in (0,\infty)$

$$
\lim_{x\to x_i^+} g_1(v(t))\frac{D\mu(t)}{D\mathcal{L}^1}(x) = c_i(v(t))\int_{\{x_i\}} d\mu(t).
$$

Theorem 2.17. *For every* $\mu_0 \in \mathcal{M}(\mathbb{R})$ *such that* $supp(\mu_0) \subset [x_0, x_N]$ *there exists a unique measure-transmission solution of problem* (1)-(3) *in the sense of Definition 2.16.*

Proof. According to Theorem 7.1 and Theorem 8.1 from [6] there exists a unique solution of problem (1)-(3) in the sense of [6, Definition 3.3] belonging to $Lip^{loc}([0,\infty),(\mathcal{M},\rho_F))$. Nevertheless, every solution constructed in [6] belongs in fact to $Lip^{loc}([0,\infty),(\mathcal{M},\rho_F))$ and is a solution of problem (1)-(3) in the sense of Definition 2.16. This is due to the fact that proof of Lemma 4.9 from [6] which states that for every $T > 0$ the mapping $t \mapsto \mu(t)$ is in $Lip([0,T],(\mathcal{M},\rho_F)$ demonstrates in fact more – that for every $T > 0$ the mapping $t \mapsto \mu(t)$ is in $Lip([0,T],(\mathcal{M},\rho_{MT})$. Existence follows. Uniqueness is due to the fact that $C([0,\infty),(\mathcal{M},\rho_{MT})) \subset C([0,\infty),(\mathcal{M},\rho_F))$ and uniqueness in the latter space. □

Modelling Remark 2.18. *Space* $C([0,\infty),\mathcal{M}(\rho_F))$ *used in [6] is inadequate for this problem. Instead we use* $C([0,\infty),\mathcal{M}(\rho_{MT}))$ *which seems to be optimal.*

3 Stability Theorem

The point of departure for the following considerations is the system of equations:

$$
\partial_t\mu(t) + \partial_x(g_1(v(t))\mathbf{1}_{x\neq x_i}(x)\mu(t)) = 0, \tag{5}
$$

$$
g_1(v(t))\frac{D\mu(t)}{D\mathcal{L}^1}(x_i^+) = c_i(v(t))\int_{\{x_i\}} d\mu(t), \qquad i = 0,\ldots,N \tag{6}
$$

$$
\mu(0) = \mu_0. \tag{7}
$$

This is a simplification of system (1)-(3) obtained by taking $p = 0$.

We use the Measure-Transmission metric to express the stability of the system with respect to perturbations of the initial condition. The rationale for this metric was explained in Section 2.1 and existence and uniqueness of solutions follows by Theorem 2.17.

We consider system (5)-(7) with two different initial measures $\mu_1(0)$ and $\mu_2(0)$. We prove that $\mu_1(t)$ stays close to $\mu_2(t)$ provided that $\mu_1(0)$ is close enough to $\mu_2(0)$. More precisely, our goal is the following

[4] Note the difference in comparison to [6]. Here we employ the Measure-Transmission metric instead of flat metric. This makes the third point from [6, Definition 3.3] obsolete and simplifies the second.

Theorem 3.1 (Stability theorem)**.** *Let $\mu_1(t)$ and $\mu_2(t)$ be two solutions of system* (5)-(7) *in the sense of Definition 2.16 corresponding to initial conditions $\mu_1(0)$ and $\mu_2(0)$, respectively. There exist constants α, β, dependent only on* $\sup(c)$, $\sup(g_1)$, $\min(g_1)$, $\mathrm{Lip}(g_1)$, $\mathrm{Lip}(c)$, $Var(\mu_1(0))$, $Var(\mu_2(0))$ *such that*

$$\rho_{MT}(\mu_1(t), \mu_2(t)) \leq e^{\alpha \left\lceil \frac{t}{\beta} \right\rceil} \rho_{MT}(\mu_1(0), \mu_2(0)),$$

where $\left\lceil \frac{t}{\beta} \right\rceil$ is the smallest integer greater or equal $\frac{t}{\beta}$ [5].

3.1 Structure of proof

We proceed in the following steps:

1. We obtain an estimate for $\int_0^T |v_1(t) - v_2(t)| dt$ in terms of $\rho_{MT}(\mu_1(0), \mu_2(0))$ and variations of $\mu_1(0)$ and $\mu_2(0)$ on a neighborhood of x_N (Nonlinear Estimate).
2. We obtain an estimate for $\rho_{MT}(\mu_1(t), \mu_2(t))$ for small t in terms of $\int_0^t |v_1(s) - v_2(s)| ds$ and $\rho_{MT}(\mu_1(0), \mu_2(0))$ (Linear Estimate).
3. We substitute the Nonlinear Estimate into the Linear Estimate to obtain an estimate of $\rho_{MT}(\mu_1(t), \mu_2(t))$ in terms of $\rho_{MT}(\mu_1(0), \mu_2(0))$ for small t.
4. We prolong the estimate for large t.

3.2 Nonlinear estimate

Our goal here is to estimate $\int_0^T |v_1(t) - v_2(t)| dt$ in terms of $\rho_{MT}(\mu_1(0), \mu_2(0))$. We begin with the implicit formula for v_j.

$$v_j(t) = \int_{[x_N - \int_0^t g_1(v_j(s)) ds, x_N]} d\mu_j(0) = \int_{[x_N - G_j(t), x_N]} d\mu_j(0),$$

where we define $G_j(t) := \int_0^T g_1(v_j(s)) ds$. Denoting $\min(G_j) := \min(G_1, G_2)$ and $\max(G_j) := \max(G_1, G_2)$ we obtain

$$\int_0^T |v_1(t) - v_2(t)| dt = \int_0^T \left| \int_{[x_N - G_1(t), x_N]} d\mu_1(0) - \int_{[x_N - G_2(t), x_N]} d\mu_2(0) \right| dt \leq$$
$$\int_0^T \left| \int_{[x_N - \min(G_j), x_N]} d(\mu_1(0) - \mu_2(0) \right| dt + \int_0^T \int_{[x_N - \max(G_j), x_N - \min(G_j))} d(\mu_1(0) + \mu_2(0)) dt = I_1 + I_2$$

Let $B := \left\{ t : \int_{[x_N - \min(G_j)(t), x_N]} d(\mu_1(0) - \mu_2(0)) \geq 0 \right\}$. Then, by Fubini theorem (see Figure 3.2), I_1 is equal to

$$\begin{aligned}
&\int_0^T (\mathbf{1}_B - \mathbf{1}_{\mathbb{R} \setminus B}) \int_{[x_N - \min(G_i), x_N]} d(\mu_1(0) - \mu_2(0)) dt \\
= &\int_{[x_N - \min(G_j(T)), x_N]} \left(\int_{\tau_{\min}(x_b)}^T (\mathbf{1}_B - \mathbf{1}_{\mathbb{R} \setminus B})(t) dt \right) d(\mu_1(0) - \mu_2(0))(x_b) \\
= &\int_{[x_N - \min(G_j(T)), x_N]} \chi(x_b) d(\mu_1(0) - \mu_2(0))(x_b).
\end{aligned}$$

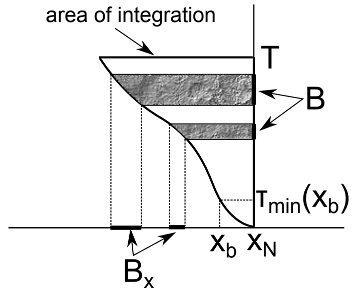

Figure 1: The area of integration in I_1. The integral can be interpreted as a double integral of function $(\mathbf{1}_B - \mathbf{1}_{\mathbb{R}\setminus B})(t)$, which is negative in the shaded region, with respect to measure $(\mu_1(0) - \mu_2(0)) \otimes dt$. B_x is the set of all x for which $\int_{[x,x_N]} d(\mu_1(0) - \mu_2(0))$ is negative.

Function $\chi(x) := \int_{\tau_{\min}(x)}^{T} (\mathbf{1}_B - \mathbf{1}_{\mathbb{R}\setminus B})(t)dt$ after prolongation by 0 to the left of $x_{\min}$ (x_{min} is by definition the unique solution of equation $\tau_{\min}(x_{\min}) = T$) and to the right of x_N belongs to $W^{1,\infty}_{MT}$. Moreover,

$$|\chi(x)| \leq T$$

and

$$|\chi'(x)| = |(\mathbf{1}_B - \mathbf{1}_{\mathbb{R}\setminus B})||\tau'_{\min}(x)| \leq \frac{1}{\min(g_1)}.$$

This allows us to conclude that

$$I_1 \leq \max\left(\frac{1}{\min(g_1)}, T\right) \rho_{MT}(\mu_1(0), \mu_2(0)).$$

Using Fubini theorem again, we estimate I_2 (compare [6, Figure 5]),

$$I_2 \leq \sup_{x\in J_{min}} |\tau_1(x) - \tau_2(x)|(\mu_1(0)(J_{\max}) + \mu_2(0)(J_{\max})) \leq (\mu_1(0)(J_{\max}) + \mu_2(0)(J_{\max}))\frac{\mathrm{Lip}(g_1)}{\min(g_1)} \int_0^T |v_1(t) - v_2(t)|dt.$$

Combining both estimates we obtain

$$\int_0^T |v_1(t) - v_2(t)|dt \leq \max\left(\frac{1}{\min(g_1)}, T\right) \frac{1}{\left(1 - \frac{\mathrm{Lip}(g_1)}{\min(g_1)}(\mu_1(0)(J_{\max}) + \mu_2(0)(J_{max}))\right)} \rho_{MT}(\mu_1(0), \mu_2(0)). \quad (8)$$

This Nonlinear Estimate makes sense provided the quantity $\mu_1(0)(J_{\max}) + \mu_2(0)(J_{\max})$ is sufficiently small. The fact that the endpoint x_N does not belong to $J_{\max}$ is crucial. This will allow us to prolong the stability estimate for arbitrary times even though the length of the time interval T related to J_{max} cannot be controlled easily.

Remark 3.2. *For $g_1 \equiv 1$ estimate (8) turns into*

$$\int_0^T |v_1(t) - v_2(t)|dt \leq \max(1, T)\, \rho_{MT}(\mu_1(0), \mu_2(0)). \quad (9)$$

[5] In particular, $\lim_{t\to 0^+} e^{\alpha\left\lceil \frac{t}{\beta}\right\rceil} = e^{\alpha}$.

3.3 Linear estimate

In this section we estimate the distance of $\mu_1(T)$ and $\mu_2(T)$, i.e. the quantity

$$\rho_{MT}(\mu_1(T), \mu_2(T)) := \sup_{\psi \in B^{1,\infty}_{MT}} \int_{\mathbb{R}} \psi d(\mu_2(T) - \mu_1(T)).$$

The main idea consists in breaking up the integral $\int_{\mathbb{R}} \psi d(\mu_2(T) - \mu_1(T))$ into parts that can be bounded in terms of $\int_0^T |v_1(s) - v_2(s)|ds$ and parts that make up the test function ψ^0, which add up to $\int_{\mathbb{R}} \psi^0 d(\mu_2(0) - \mu_1(0))$. Both will be then estimated by $C_1(t)\rho_{MT}(\mu_1(0), \mu_2(0))$.

To achieve this goal, we assume without loss of generality (compare Remark 3.6) that $G_1(T) \leq G_2(T)$ and split the integral into three main components with respect to characteristics:

- Characteristics starting in $(x_{i-1}, x_i - G_2(T))$ (Fig. 2) – terms I_i,
- Characteristics starting in $[x_i - G_2(T), x_i - G_1(T))$ (Fig. 3) – terms T_i,
- Characteristics starting in $[x_i - G_1(T), x_i]$ (Fig. 4) – terms D_i.

We obtain:

$$\int_{\mathbb{R}} \psi d(\mu_2(T) - \mu_1(T)) = D_0 + (I_1 + T_1 + D_1) + (I_2 + T_2 + D_2) + \cdots + (I_N + T_N + D_N)$$

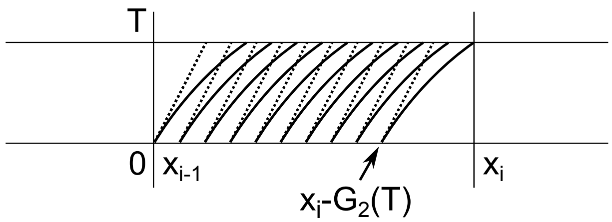

Figure 2: Characteristics generated by $g_1(v_1)$ (dotted) and $g_1(v_2)$ (solid) starting in $(x_{i-1}, x_i - G_2(T))$

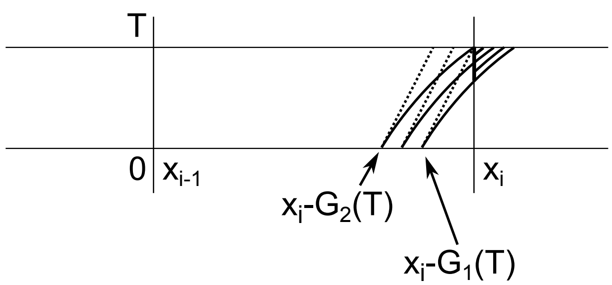

Figure 3: Characteristics generated by $g_1(v_1)$ (dotted) and $g_1(v_2)$ (solid) starting in $[x_i - G_2(T), x_i - G_1(T))$. Characteristics corresponding to $g_1(v_2)$ arrive in x_i before time T and generate fans of characteristics whereas those corresponding to $g_1(v_1)$ do not.

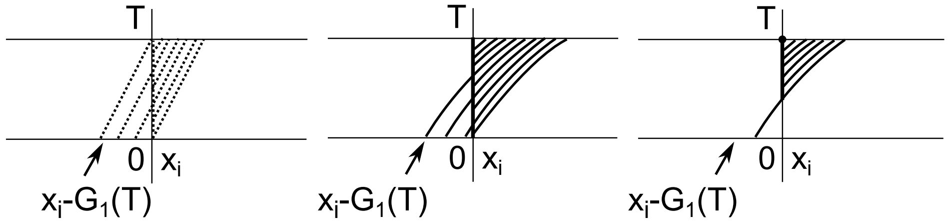

Figure 4: Characteristics generated by $g_1(v_1)$ (dotted, left panel) and $g_1(v_2)$ (solid, middle panel) starting in interval $[x_i - G_1(T), x_i]$. After arriving in x_i a given characteristic either spends an arbitrary period of time in x_i before leaving x_i or stays there until time T. Thus every characteristic coming to x_i branches generating a fan of characteristics (right panel)

This decomposition of solution into atoms is similar in spirit to the idea of superposition solution developed for linear transport equation by Ambrosio, Crippa and Maniglia, (see e.g. [1, 12, 13]) and follows by the definition of solutions ([6, Section 4]) [6]. In the following estimates we assume $i = 1$. The terms for other i are estimated likewise.

3.3.1 Characteristics starting in $(x_0, x_1 - G_2(T))$

$$\begin{aligned} I_1 = \int_{(x_0+G_2(T),x_1)} \psi d\mu_2(T) - \int_{(x_0+G_1(T),x_1-G_2(T)+G_1(T))ds} \psi d\mu_1(T) &= \\ \int_{(x_0,x_1-G_2(T))} \left[\psi\left(x_b + G_2(T)\right) d\mu_2(0)(x_b) - \psi\left(x_b + G_1(T)\right) d\mu_1(0)(x_b)\right] &= \\ \int_{(x_0,x_1-G_2(T)} \psi\left(x_b + G_1(T)\right) d(\mu_2(0) - \mu_1(0)) + \int_{(x_0,x_1-G_2(T)} \left[\psi\left(x_b + G_2(T)\right) - \psi\left(x_b + G_1(T)\right)\right] d\mu_2(0) &= \\ U^{I_1} + V^{I_1} & . \end{aligned}$$

3.3.2 Characteristics starting in $[x_1 - G_2(T), x_1 - G_1(T))$

Note that for μ_1 these characteristics are simple curves without ramifications like in section 3.3.1. In case of μ_2, however, they reach x_1 before time T and therefore may branch. We obtain

$$T_1 = \int_{[x_1-G_2(T),x_1-G_1(T))} \left\{ \psi(x_1) e^{-\int_{\tau_2(x_b)}^T c_1(v_2(s))ds} d\mu_2(0)(x_b) + \right.$$
$$\left. \left(\int_{\tau_2(x_b)}^T e^{-\int_{\tau_2(x_b)}^r c_1(v_2(s))ds} c_1(v_2(r)) \psi\left(x_1 + \int_r^T g_1(v_1(s))ds\right) dr \right) d\mu_2(0)(x_b) - \psi\left(x_b + G_1(T)\right) d\mu_1(0)(x_b) \right\}$$

[6] More precisely, this follows by the fact that in conservative case at hand ($p = 0$) the measure is transported along (branching) characteristics ([6, Section 4, Formulas 18,19,20]). Branching is expressed quantitatively by functions f_i defined by ([6, Section 4, Formula 17]). These functions correspond to coefficients $c_i(v(r))e^{-\int_{t_1}^{t_2} c_i(v(s))ds}$.

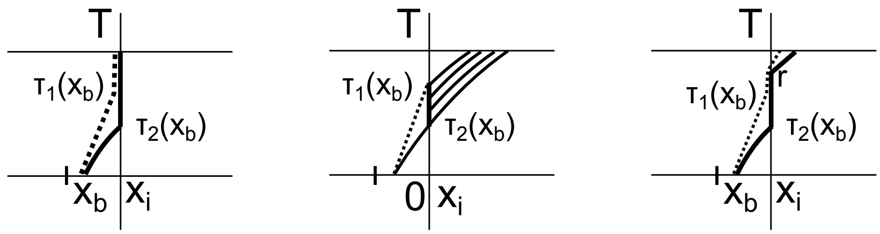

Figure 5: Sample characteristics starting in $x_b \in [x_1 - G_1(T), x_1]$. Left panel. Characteristics starting in x_b and both ending in x_i. Middle panel. The fan of characteristics arriving at time $\tau_2(x_b)$ and leaving before $\tau_1(x_b)$ is small provided $|\tau_1(x_b) - \tau_2(x_b)|$ is small. Right panel. Characteristics starting in x_b and *both* branching off at time r.

Consecutive terms in the integrand correspond to characteristics related to $\mu_2(0)$ ending in x_1, related to $\mu_2(0)$ ending in (x_1, x_2) and related to $\mu_1(0)$. Further calculations lead to

$$
\begin{aligned}
T_1 = \int_{[x_1-G_2(T),x_1-G_1(T))} \psi\left(x_b + G_1(T)\right) d(\mu_2(0) - \mu_1(0)) \quad + \\
\int_{[x_1-G_2(T),x_1-G_1(T))} \left(\psi(x_1) - \psi\left(x_b + G_1(T)\right)\right) d\mu_2(0) \quad + \\
\int_{[x_1-G_2(T),x_1-G_1(T))} \psi(x_1)\left(e^{-\int_{\tau_2(x_b)}^T c_1(v_2(s))ds} - 1\right) d\mu_2(0) \quad + \\
\int_{[x_1-G_2(T),x_1-G_1(T))} \left[\int_{\tau_2(x_b)}^T c_1(v_2(r)) e^{-\int_{\tau_2(x_b)}^r c_1(v_2(s))ds} \psi\left(x_1 + \int_r^T g_1(v_2(s))ds\right) dr\right] d\mu_2(0) \quad = \\
U^{T_1} + V_1^{T_1} + V_2^{T_1} + V_3^{T_1}
\end{aligned}
$$

3.3.3 Characteristics starting in $[x_1 - G_1(T), x_1]$

We subdivide the characteristics into three groups (See Fig. 5):

- those ending in x_1,
- those ending in (x_1, x_2) and branching off between $\tau_2(x_b)$ and $\tau_1(x_b)$,
- those ending in (x_1, x_2) and branching off between $\tau_1(x_b)$ and T .

This leads to:

$$
\begin{aligned}
D_1 = \int_{[x_1-G_1(T),x_1]} \psi(x_1) e^{-\int_{\tau_2(x_b)}^T c_1(v_2(s))ds} d\mu_2(0) - \int_{[x_1-G_1(T),x_1]} \psi(x_1) e^{-\int_{\tau_1(x_b)}^T c_1(v_1(s))ds} d\mu_1(0) + \\
\int_{[x_1-G_1(T),x_1]} \left(\int_{\tau_2(x_b)}^{\tau_1(x_b)} c_1(v_2(r)) e^{-\int_{\tau_2(x_b)}^r c_1(v_2(s))ds} \psi\left(x_1 + \int_r^T g_1(v_2(s))ds\right) dr\right) d\mu_2(0) + \\
\int_{[x_1-G_1(T),x_1]} \int_{\tau_1(x_b)}^T \Bigg[c_1(v_2(r)) e^{-\int_{\tau_2(x_b)}^r c_1(v_2(s))ds} \psi\left(x_1 + \int_r^T g_1(v_2(s))ds\right) dr d\mu_2(0) - \\
c_1(v_1(r)) e^{-\int_{\tau_1(x_b)}^r c_1(v_1(s))ds} \psi\left(x_1 + \int_r^T g_1(v_1(s))ds\right) dr d\mu_1(0) \Bigg] = \\
1^\circ + 2^\circ + 3^\circ
\end{aligned}
$$

$$
\begin{aligned}
1^\circ &= \psi(x_1)\int_{[x_1-G_1(T),x_1]} e^{-\int_{\tau_1(x_b)}^T c_1(v_1(s))ds} d(\mu_2(0)-\mu_1(0)) + \\
&\quad \psi(x_1)\int_{[x_1-G_1(T),x_1]} \left(e^{-\int_{\tau_2(x_b)}^T c_1(v_2(s))ds} - e^{-\int_{\tau_1(x_b)}^T c_1(v_1(s))ds}\right) d\mu_2(0) = U_1^{D_1} + V_1^{D_1}, \\
2^\circ &= V_2^{D_1} \\
3^\circ &= \int_{[x_1-G_1(T),x_1]} \left(\int_{\tau_1(x_b)}^T c_1(v_1(r)) e^{-\int_{\tau_1(x_b)}^r c_1(v_1(s))ds} \psi\left(x_1 + \int_r^T g_1(v_1(s))ds\right) dr\right) d(\mu_2(0)-\mu_1(0)) + \\
&\quad \int_{[x_1-G_1(T),x_1]} \int_{\tau_1(x_b)}^T \Bigg[c_1(v_2(r)) e^{-\int_{\tau_2(x_b)}^r c_1(v_2(s))ds} \psi\left(x_1 + \int_r^T g_1(v_2(s))ds\right) - \\
&\quad c_1(v_1(r)) e^{-\int_{\tau_1(x_b)}^r c_1(v_1(s))ds} \psi\left(x_1 + \int_r^T g_1(v_1(s))ds\right) \Bigg] dr d\mu_2(0) = U_2^{D_1} + V_3^{D_1}
\end{aligned}
$$

Collecting similar terms we obtain:

$$
I_1 + T_1 + D_1 = \left(U^{I_1} + U^{T_1} + U^{D_1} + U^{D_2}\right) + \left(V^{I_1} + V_1^{T_1} + V_2^{T_1} + V_3^{T_1} + V_1^{D_1} + V_2^{D_1} + V_3^{D_1}\right)
$$

3.3.4 U terms

$$
\left(U^{I_1} + U^{T_1} + U^{D_1} + U^{D_2}\right) = \int_{(x_0,x_1]} \psi^0(x_b) d(\mu_2(0)-\mu_1(0))(x_b),
$$

where

$$
\psi^0(x_b) = \begin{cases} \psi(x_b + G_1(T)) \\ \qquad \text{for } x_0 < x_b < x_1 - G_1(T) \\ \psi(x_1) e^{-\int_{\tau_1(x_b)}^T c_1(v_1(s))ds} + \int_{\tau_1(x_b)}^T c_1(v_1(r)) e^{-\int_{\tau_1(x_b)}^r c_1(v_1(s))ds} \psi\left(x_1 + \int_r^T g_1(v_1(s))ds\right) dr \\ \qquad \text{for } x_1 - G_1(T) \le x_b \le x_1 \end{cases}
$$

Note that ψ^0 is continuous in $x_1 - G_1(T)$ and right-continuous in x_1. Let us compute explicitly the derivative of ψ^0 for $x_1 - G_1(T) < x_b < x_1$.

$$
\begin{aligned}
(\psi^0)'(x_b) = \tau_1'(x_b) c_1(v_1(\tau_1(x_b))) \psi(x_1) e^{-\int_{\tau_1(x_b)}^T c_1(v_1(s))ds} - \tau_1'(x_b) c_1(v_1(\tau_1(x_b))) \psi\left(x_1 + \int_{\tau_1(x_b)}^T g_1(v_1(s))ds\right) + \\
\int_{\tau_1(x_b)}^T c_1(v_1(r)) e^{-\int_{\tau_1(x_b)}^r c_1(v_1(s))ds} \tau_1'(x_b) c_1(v_1(\tau_1(x_b))) \psi\left(x_1 + \int_r^T g_1(v_1(s))ds\right) dr.
\end{aligned}
$$

Using $\tau_1'(x_b) \le \frac{1}{\min(g_1)}$, we arrive at

$$
|(\psi^0)'(x_b)| \le \frac{1}{\min(g_1)} \left(\sup(\psi)\sup(c_1) + \sup(\psi)\sup(c_1) + \sup(\psi) T (\sup(c_1))^2\right). \tag{10}
$$

Thus,

$$
\begin{aligned}
|\psi^0(x_b)| &\le \sup(\psi) \\
\left|\frac{d}{dx_b}\psi^0(x_b)\right| &\le \max\left(\sup(\psi), \frac{\sup(\psi')\sup(c_1)}{\min(g_1)}(2 + T\sup(c))\right)
\end{aligned}
$$

These estimates hold for all $x_b \in [x_0, x_N]$.

3.3.5 V terms

$$
\begin{aligned}
|V^{I_1}| &\leq \operatorname{Lip}(\psi)\operatorname{Lip}(g_1)\mu_2(0)((x_0, x_1 - G_2(T)) \int_0^T |v_2(s) - v_1(s)| ds \\
|V_1^{T_1}| &\leq \operatorname{Lip}(\psi)\operatorname{Lip}(g_1)\mu_2(0)([x_1 - G_2(T), x_1 - G_1(T)) \int_0^T |v_2(s) - v_1(s)| ds \\
|V_2^{T_1}| &\leq |\psi(x_1)| \sup(c_1) \sup(T - \tau_2(x_b))\mu_2(0)([x_1 - G_2(T), x_1 - G_1(T)) \\
&\leq |\psi(x_1)| \sup(c_1) \frac{\operatorname{Lip}(g_1)}{\min(g_1)} \mu_2(0)([x_1 - G_2(T), x_1 - G_1(T)) \int_0^T |v_2(s) - v_1(s)| ds \\
|V_3^{T_1}| &\leq \sup(\psi) \sup(c_1) \frac{\operatorname{Lip}(g_1)}{\min(g_1)} \mu_2(0)([x_1 - G_2(T), x_1 - G_1(T)) \int_0^T |v_2(s) - v_1(s)| ds \\
|V_1^{D_1}| &\leq |\psi(x_1)| \left| \int_{\tau_2(x_b)}^T c_1(v_2(s) ds - \int_{\tau_1(x_b)}^T c_1(v_1(s)) ds \right| \mu_2(0)([x_1 - G_1(T), x_1]) \\
&\leq |\psi(x_1)| \left(\operatorname{Lip}(c_1) \int_0^T |v_2(s) - v_1(s)| ds + \sup(c_1) \sup |\tau_2(x_b) - \tau_1(x_b)| \right) \mu_2(0)([x_1 - G_1(T), x_1]) \\
&\leq |\psi(x_1)| \left(\operatorname{Lip}(c_1) + \sup(c_1) \frac{\operatorname{Lip}(g_1)}{\min(g_1)} \right) \mu_2(0)([x_1 - G_1(T), x_1]) \int_0^T |v_2(s) - v_1(s)| ds, \\
|V_2^{D_1}| &\leq \sup_{x_b} |\tau_2(x_b) - \tau_1(x_b)| \sup(c_1) \sup(\psi) \mu_2(0)([x_1 - G_1(T), x_1]) \\
&\leq \frac{\operatorname{Lip}(g_1)}{\min(g_1)} \sup(c_1) \sup(\psi) \mu_2(0)([x_1 - G_1(T), x_1]) \int_0^T |v_1(s) - v_2(s)| ds
\end{aligned}
$$

To estimate $V_3^{D_1}$ let us first consider the inner integral

$$
\begin{aligned}
&\int_{\tau_1(x_b)}^T \Bigg[c_1(v_2(r)) e^{-\int_{\tau_2(x_b)}^r c_1(v_2(s)) ds} \psi \left(x_1 + \int_r^T g_1(v_2(s)) ds \right) - \\
&\qquad c_1(v_1(r)) e^{-\int_{\tau_1(x_b))}^r c_1(v_1(s)) ds} \psi \left(x_1 + \int_r^T g_1(v_1(s)) ds \right) \Bigg] dr = \\
&\int_{\tau_1(x_b)}^T (c_1(v_2(r)) - c_1(v_1(r)) e^{-\int_{\tau_2(x_b)}^r c_1(v_2(s)) ds} \psi \left(x_1 + \int_r^T g_1(v_2(s)) ds \right) dr + \\
&\int_{\tau_1(x_b)}^T c_1(v_1(r)) \left(e^{-\int_{\tau_2(x_b)}^r c_1(v_2(s)) ds} - e^{-\int_{\tau_1(x_b)}^r c_1(v_1(s)) ds} \right) \psi \left(x_1 + \int_r^T g_1(v_2(s)) ds \right) dr + \\
&\int_{\tau_1(x_b)}^T c_1(v_1(r)) e^{-\int_{\tau_1(x_b)}^r c_1(v_1(s)) ds} \left(\psi \left(x_1 + \int_r^T g_1(v_2(s)) ds \right) - \psi \left(x_1 + \int_r^T g_1(v_1(s)) ds \right) \right) dr = \\
&I_\alpha + I_\beta + I_\gamma
\end{aligned}
$$

$$
\begin{aligned}
|I_\alpha| &\leq \sup(\psi) \operatorname{Lip}(c_1) \int_0^T |v_2(s) - v_1(s))| ds \\
|I_\beta| &\leq T \sup(\psi) \sup(c_1) \left(\sup(c_1) \frac{\operatorname{Lip}(g_1)}{\min(g_1)} + \operatorname{Lip}(c_1) \right) \int_0^T |v_2(s) - v_1(s)| ds \\
|I_\gamma| &\leq T \sup(c_1) \operatorname{Lip}(\psi) \operatorname{Lip}(g_1) \int_0^T |v_2(s) - v_1(s)| ds
\end{aligned}
$$

Thus,

$$\begin{aligned}
|V_3^{D_1}| &\leq (|I_\alpha| + |I_\beta| + |I_\gamma|)\mu_2(0)([x_1 - G_1(T), x_1]) \\
&\leq \left(\sup(\psi)\mathrm{Lip}(c_1) + \sup(\psi)\sup(c_1)T\left(\sup(c_1)\frac{\mathrm{Lip}(g_1)}{\min(g_1)} + \mathrm{Lip}(c_1)\right) + \mathrm{Lip}(\psi)\sup(c_1)T\mathrm{Lip}(g_1)\right) \\
&\quad \mu_2(0)([x_1 - G_1(T), x_1])\int_0^T |v_1(s) - v_2(s)|ds.
\end{aligned}$$

Combining U-terms and V-terms for $i \in \{0, \ldots, N\}$ we obtain

$$\left|\int_{\mathbb{R}} \psi d(\mu_2(T) - \mu_1(T))\right| \leq |D_0| + |I_1 + T_1 + D_1| + |I_2 + T_2 + D_2| + \cdots + |I_N + T_N + D_N| \leq \int_{\mathbb{R}} \psi^0 d(\mu_2(0) - \mu_1(0)) + \left\{\mathrm{Lip}(\psi)\mathrm{Lip}(g_1) + 2\sup(\psi)\sup(c)\frac{\mathrm{Lip}(g_1)}{\min(g_1)} + \sup(\psi)\mathrm{Lip}(c)T\sup(c)\left[\sup(\psi)\left(\sup(c)\frac{\mathrm{Lip}(g_1)}{\min(g_1)} + \mathrm{Lip}(c)\right) + \mathrm{Lip}(\psi)\mathrm{Lip}(g_1)\right]\right\} Var(\mu_2(0))\int_0^T |v_2(s) - v_1(s)|ds$$

Taking into account that $\|\psi^0\|_{W^{1,\infty}_{MT}} \leq \|\psi\|_{W^{1,\infty}_{MT}} \max\left(1, \frac{\sup(c)}{\min(g_1)}(2 + T\sup(c))\right)$ we obtain

$$\rho_{MT}(\mu_1(T), \mu_2(T)) = \sup_{\psi \in B^{1,\infty}_{MT}} \int_{\mathbb{R}} \psi d(\mu_2(T) - \mu_1(T)) \leq \max\left(1, \frac{\sup(c)}{\min(g_1)}(2 + T\sup(c))\right)\rho_{MT}(\mu_1(0), \mu_2(0)) + \left\{\mathrm{Lip}(g_1) + 2\sup(c)\frac{\mathrm{Lip}(g_1)}{\min(g_1)} + \mathrm{Lip}(c) + T\sup(c)\left[\sup(c)\frac{\mathrm{Lip}(g_1)}{\min(g_1)} + \mathrm{Lip}(c) + \mathrm{Lip}(g_1)\right]\right\} Var(\mu_2(0))\int_0^T |v_2(s) - v_1(s)|ds.$$

This in combination with (8) leads to the

Corollary 3.3 (**Local in time stability estimate**).

$$\rho_{MT}(\mu_1(T), \mu_2(T)) \leq C_1(T)\rho_{MT}(\mu_1(0), \mu_2(0)), \tag{11}$$

where

$$\begin{aligned}
C_1(T) &= \max\left(1, \frac{\sup(c)}{\min(g_1)}(2 + T\sup(c))\right) + \\
&\quad \left\{\mathrm{Lip}(g_1) + 2\sup(c)\frac{\mathrm{Lip}(g_1)}{\min(g_1)} + \mathrm{Lip}(c) + T\sup(c)\left[\sup(c)\frac{\mathrm{Lip}(g_1)}{\min(g_1)} + \mathrm{Lip}(c) + \mathrm{Lip}(g_1)\right]\right\} \times \\
&\quad \times Var(\mu_2(0))\max\left(\frac{1}{\min(g_1)}, T\right)\frac{1}{\left(1 - \frac{\mathrm{Lip}(g_1)}{\min(g_1)}(\mu_1(0)(J_{\max}) + \mu_2(0)(J_{max}))\right)}.
\end{aligned} \tag{12}$$

Example 3.4. *Take $\mu(0) = \delta_{x_1}$ and $\mu^\epsilon(0) = \delta_{x_1-\epsilon}$ as well as $g_1 \equiv 1$ and c_1 constant. Then,*

$$\rho_{MT}(\mu(0), \mu^\epsilon(0)) = \epsilon.$$

Let these measures evolve according to equation (5). For $t = \epsilon$ we obtain

$$\begin{aligned}\mu^{\epsilon}(t=\epsilon) &= \delta_{x_1}, \\ \mu(t=\epsilon) &= e^{-c_1\epsilon}\delta_{x_1} + c_1 e^{-c_1(\epsilon - x)}\mathbf{1}_{[0,\epsilon]}(x)\mathcal{L}^1(dx).\end{aligned}$$

Hence,

$$\rho_{MT}(\mu(\epsilon),\mu^{\epsilon}(\epsilon)) = 2\left(1 - e^{-c_1\epsilon}\right) \simeq 2c_1\epsilon = 2c_1\rho_{MT}(\mu(0),\mu^{\epsilon}(0)).$$

In other words, for every ϵ there exists a pair of measures $\mu_1(0)$ and $\mu_2(0)$ for which $\rho_{MT}(\mu_1(\epsilon),\mu_2(\epsilon)) \simeq 2c_1\rho_{MT}(\mu_1(0),\mu_2(0))$.

Example 3.5. *Take two initial measures:*

$$\mu(0) = \delta_{x_N} + \delta_y,$$

$$\mu^{\epsilon}(0) = \delta_{x_N - \epsilon} + \delta_y,$$

where $y \in (x_{N-1}, x_N)$ is such that $|x_N - y| > 1$. Clearly, $\rho_{MT}(\mu(0),\mu^{\epsilon}(0)) = \epsilon$. Take

$$g_1(v) = \begin{cases} \underline{g} & \textit{for } v = 0, \\ 1 & \textit{for } v = 1. \end{cases}$$

Let the measures evolve according to equation (5). For $\tilde{t} = \frac{\epsilon}{\underline{g}}$ we obtain $\mu(\tilde{t}) = \delta_{x_N} + \delta_{y+1}$ and $\mu^{\epsilon}(\tilde{t}) = \delta_{x_N} + \delta_{y+\epsilon/\underline{g}}$. Thus,

$$\rho_{MT}\left(\mu\left(\frac{\epsilon}{\underline{g}}\right), \mu^{\epsilon}\left(\frac{\epsilon}{\underline{g}}\right)\right) = \epsilon\left(\frac{1}{\underline{g}} - 1\right) = \left(\frac{1}{\underline{g}} - 1\right)\rho_{MT}(\mu_1(0),\mu_2(0)).$$

Letting $\epsilon \to 0$ leads us to conclusion that $C_1(0^+) \geq \left(\frac{1}{\underline{g}} - 1\right)$.

These two examples show that it is impossible to obtain an estimate with $C_1(0^+) = 1$ for arbitrary data.

Remark 3.6. *It may happen that characteristics generated by $g_1(v_1)$ and $g_1(v_2)$ cross in such a way that although $G_1(T) < G_2(T)$ there exist certain x_b for which $\tau_1(x_b) < \tau_2(x_b)$. The reader will easily modify the proof to encompass such behavior.*

3.4 Stability estimate for large times

Our goal is to obtain a global in time stability estimate with constant which depends only on the total mass of measures $\mu_1(0), \mu_2(0)$ and not on the initial mass distribution, i.e. the detailed structure of initial measures. We shall iterate estimate (11). Since $C_1(T)$ given by formula (12) is strongly dependent on the values $\mu_1(0)(J_{max})$ and $\mu_2(0)(J_{max})$ we cannot use a constant time step. Instead we choose such time increments, for which mass contained in J_{max} is small enough for (11) to make sense and large enough to be able to prolong the estimate in a 'reasonable' number of steps. More precisely, we choose time points $0 = T_0 < T_1 < T_2 < \ldots$ in such a way that for $j \in \{1, 2\}, k = 0, 1, \ldots$

$$\mu_j(T_k)(J_{max}) \leq L := \frac{1}{4}\frac{\min(g_1)}{\mathrm{Lip}(g_1)}, \tag{13}$$

$$\Delta T_k := T_{k+1} - T_k \leq \min\left(1, \min_{i\in\{0,\ldots,N-1\}} \frac{|x_i - x_{i+1}|}{\sup(g_1)}\right) = \min(1, T_{max}) \tag{14}$$

and ΔT_k are maximal. To this end, we first choose a maximum ΔT_0 for which both (13) and (14) hold. Next, we take a maximum ΔT_1, for which both (13) and (14) hold and so on. Repeating this procedure, we observe that

Proposition 3.7. *Either*

$$\Delta T_k = \min(1, T_{max})$$

or

$$\mu_1(T_k)(J^{lc}_{max}) + \mu_2(T_k)(J^{lc}_{max}) \geq L,$$

where lc stands for left closure of an interval, i.e. $(a,b)^{lc} = [a,b)$.

Proof. If $\Delta T < \min(1, T_{max})$ then either for $\mu_1(0)(J_{max}) = L$ or $\mu_2(T_k)(J_{max}) = L$ or both $\mu_1(T_k)(J_{max}) < L$ and $\mu_2(T_k)(J_{max})) < L$. In the latter case either $\mu_1(T_k)(J^{lc}_{max}) > L$ or $\mu_2(T_k)(J^{lc}_{max}) > L$ due to the fact that ΔT_k is the maximum time interval for which (13) holds. □

Corollary 3.8. *Let* $T_{intmin} := \frac{|x_N - x_{N-1}|}{\min(g_1)}$[7]. *It takes at most*

$$It_1 = \frac{T_{intmin}}{\min(1, T_{max})} + \left\lceil \frac{\mu_1(0)((x_{N-1}, x_N)) + \mu_2(0)((x_{N-1}, x_N))}{L} \right\rceil$$

time steps (iterations) defined above for the mass related to at least one of the measures $\mu_1(0), \mu_2(0)$ *from interval* (x_{N-1}, x_N) *to arrive* [8] *in* x_N.

Proof. Every iteration, according to Proposition 3.7, falls into one of two categories: by time $\min(1, T_{max})$ or by mass at least L. The number of iterations of the former type is bounded by $\frac{T_{intmin}}{\min(1,T_{max})}$ and of the latter by $\left\lceil \frac{\mu_1(0)((x_{N-1},x_N)) + \mu_2(0)((x_{N-1},x_N))}{L} \right\rceil$ □

To obtain a global in time constant, we observe that due to the fact that for all $t > 0$ $Var(\mu_j(t)) = Var(\mu_j(0))$ we have

$$It_1 \leq It_2 := \frac{T_{intmin}}{\min(1, T_{max})} + \left\lceil \frac{Var(\mu_1(0) + Var(\mu_2(0))}{L} \right\rceil.$$

Now, we calculate the arrival time in a different way. The mass from interval (x_{N-1}, x_N) related to either of the measures $\mu_1(0), \mu_2(0)$ requires at least time $T_{int} = \frac{|x_{N-1} - x_N|}{\sup(g_1)}$ to arrive in x_N. Thus, advancing of the above described procedure by time T_{int} requires no more than It_2 iterations. Hence, advancing by time t requires no more than $It_2 \left\lceil \frac{t}{T_{int}} \right\rceil$ iterations.

Corollary 3.9.

$$\rho_{MT}(\mu_1(t), \mu_2(t)) \leq \kappa^{\left(It_2 \left\lceil \frac{t}{T_{int}} \right\rceil\right)} \rho_{MT}(\mu_1(0), \mu_2(0)), \tag{15}$$

[7] T_{intmin} is the maximum time necessary for all characteristics starting from interval $(x_{N-1}, x_N]$ to arrive in x_N

[8] The mass flows along characteristics in one direction, i.e. forward. After arriving in x_N a piece of mass cannot leave anymore so the arrival time is uniquely defined.

where

$$\begin{aligned}
T_{int} &= \frac{|x_N - x_{N-1}|}{\sup(g_1)} \\
It_2 &= \left\lceil \frac{Var(\mu_1(0) + Var(\mu_2(0))}{L} \right\rceil \\
L &= \frac{1}{4}\frac{\min(g_1)}{\mathrm{Lip}(g_1)} \\
\kappa &= \max\left(1, \frac{\sup(c)}{\min(g_1)}(2+\sup(c))\right) + \\
& 2\left\{\mathrm{Lip}(g_1) + 2\sup(c)\frac{\mathrm{Lip}(g_1)}{\min(g_1)} + \mathrm{Lip}(c) + \sup(c)\left[\sup(c)\frac{\mathrm{Lip}(g_1)}{\min(g_1)} + \mathrm{Lip}(c) + \mathrm{Lip}(g_1)\right]\right\} \times \\
& \times Var(\mu_2(0)) \max\left(\frac{1}{\min(g_1)}, 1\right).
\end{aligned}$$

This proves Theorem 3.1.

Corollary 3.10. *For $T < T_{int}$ we have*

$$\int_0^T |v_1(t) - v_2(t)|dt \le 2\max\left(\frac{1}{\min(g_1)}, T_{int}\right)(It_1)\kappa^{It_2}\rho_{MT}(\mu_1(0), \mu_2(0))$$

Proof. By (8) and (15) we obtain

$$\int_{T_k}^{T_{k+1}} |v_1(t) - v_2(t)|dt \le$$

$$2\max\left(\frac{1}{\min(g_1)}, T_{int}\right)\rho_{MT}(\mu_1(T_k), \mu_2(T_k)) \le 2\max\left(\frac{1}{\min(g_1)}, T_{int}\right)\kappa^{It_2}\rho_{MT}(\mu_1(0), \mu_2(0)).$$

Summing from $k = 0$ to $k = N - 1$, we conclude. □

Remark 3.11. System (5)-(7) is conservative. As a result, for given $Var(\mu_1(0)), Var(\mu_2(0))$ both v_1 and v_2 are bounded. By continuity of g_1 we obtain that also $g_1(v_1)$ and $g_1(v_2)$ are bounded from below and therefore $\min(g_1) > 0$.

Remark 3.12. Time steps in iterations which lead to the global stability estimate (15) are *different* for every pair of initial measures. This is due to the fact that it is constant 'mass step' that is used rather than constant time step (see Proposition 3.7). In the end, however, the estimate has the same form for *every* pair of initial measures and depends only on their total variations. This is due to the fact that there is only finite potential for small time steps which depends only on the total variation of measures (see Corollary 3.8).

4 Summary and Applications

We proved the stability of solutions of system (5)-(7) with respect to perturbation of initial condition. The result is important both from the modelling (every reasonable model of reality has to be stable) and simulation point of view. Namely, theorem 3.1 can be applied for constructing a stable and convergent numerical scheme. The fact that $C_1(0^+) > 1$ does not play a role, since the effective algorithm is constructed along the construction of Nonlinear Estimate from Section 3.4 and *not* by direct time discretization. This is a topic for further work.

Analitically, we investigated the structure of metrics on $\mathbb{R}$ and demonstrated the method of fine-tuning the metric to the model at hand. We introduced a method of prolonging estimates which are very singular for large times.

Acknowledgement This work was supported by International Ph.D. Projects Programme of Foundation for Polish Science operated within the Innovative Economy Operational Programme 2007-2013 funded by European Regional Development Fund (Ph.D. Programme: Mathematical Methods in Natural Sciences).

References

[1] Ambrosio L, Crippa G. *Existence, Uniqueness, Stability and Differentiability Properties of the Flow Associated to Weakly Differentiable Vector Fields.* Lecture Notes of the Unione Matematica Italiana, Springer; 2008, 5: 3–57.

[2] Ambrosio L, Gigli N, Savaré G. *Gradient flows in metric spaces and in the space of probability measures.* ETH Lecture Notes in Mathematics; Birkhäuser; 2005

[3] Carrillo JA, Colombo RM, Gwiazda P, Ulikowska A. Structured Populations, Cell Growth and Measure Valued Balance Laws. *Journal of Differential Equations.* 2012; 252(4): 3245-3277.

[4] Dafermos CM. *Hyperbolic conservation laws in continuum physics.* Springer-Verlag; 2000,2005,2010.

[5] Evans LC, Gariepy RF. *Measure theory and fine properties of functions.* CRC Press; 1992.

[6] Gwiazda P, Jamróz G, Marciniak-Czochra A. Models of discrete and continuous cell differentiation in the framework of transport equation. *SIAM Journal on Mathematical Analysis* 2012; 44: 1103-1133. Preprint available at http://mmns.mimuw.edu.pl/preprints/2011-007.pdf.

[7] Gwiazda P, Lorenz T, Marciniak-Czochra A., A nonlinear structured population model: Lipschitz continuity of measure valued solutions with respect to model ingredients. *Journal of Differential Equations* 2010; 248:2703–2735.

[8] Gwiazda P, Marciniak-Czochra A, Structured population models in metric spaces. *Journal of Hyperbolic Differential Equations* 2010; 7:733–773.

[9] Knisely K, Gilbert SF. *Developmental Biology (8th ed.).* Sunderland, Mass: Sinauer Associates; 2009.

[10] Kriegstein A, Alvarez-Buylla A. The Glial Nature of Embryonic and Adult Neural Stem Cells. *Annual Review of Neuroscience* 2009; 32: 149–184.

[11] Belair J, Mackey MC, Mahaffy J. Age-structured and two-delay models for erythropoiesis, Mathematical Biosciences 1995; 128: 317–346.

[12] Maniglia S. *Fine properties and well-posedness of solutions of transport equations and of the 1st order associated ODE* PhD Thesis. Universita di Pisa; 2006.

[13] Maniglia S. Probabilistic representation and uniqueness results for measure-valued solutions of transport equations *Journal de Mathématiques Pures et Appliquées* 2007; 87: 601-626.

[14] Marciniak-Czochra A, Stiehl T, Ho AD, Jaeger W, Wagner W. Modeling asymmetric cell division in hematopoietic stem cells - regulation of self-renewal is essential for efficient repopulation. *Stem Cells and Development* 2008; 17: 1-10.

[15] Roberts K, Raff M, Alberts B, Walter P, Lewis J, Johnson A. *Molecular Biology of the Cell. 4th edition.* Garland Science: New York; 2002.

[16] Villani C, *Topics in optimal transportation. Graduate Studies in Mathematics, 58.* American Mathematical Society: Providence, RI; 2003

[17] Yakovlev AY, Mayer-Proschel M, Noble M. A stochastic model of brain cell differentiation in tissue culture. *Journal of Mathematical Biology* 1998; 37(1):49-60.