



**HAL**  
open science

# Ecological barriers and convergence: a note on geometry in spatial growth models

Giorgio Fabbri

► **To cite this version:**

Giorgio Fabbri. Ecological barriers and convergence: a note on geometry in spatial growth models. 2015. hal-01159253

**HAL Id: hal-01159253**

**<https://hal.science/hal-01159253>**

Preprint submitted on 3 Jun 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# ECOLOGICAL BARRIERS AND CONVERGENCE: A NOTE ON GEOMETRY IN SPATIAL GROWTH MODELS

GIORGIO FABBRI

**ABSTRACT.** We introduce an AK spatial growth model with a general geographical structure. The dynamics of the economy is described by a partial differential equation on a Riemannian manifold. The morphology interacts with the spatial dynamics of the capital and is one determinant of the qualitative behavior of the economy. We characterize the conditions on the geographical structure that guarantee convergence of the detrended capital across locations in the long run, and those inducing spatial capital agglomeration.

**Key words:** Dynamical spatial model; growth; agglomeration; convergence; infinite dimensional optimal control problems; Riemannian manifolds.

**Journal of Economic Literature Classification:** R1; O4; C61.

## 1. INTRODUCTION

Even if the theoretical importance of geography in development processes was already clear more than two centuries ago (see for example Smith, 1776, Book 1, Chapter 3), the effort to merge the continuous spatial dimension with benchmark growth theory models is rather recent.

In a seminal paper, Brito (2004) first<sup>1</sup> introduced spatial capital accumulation and capital mobility in the Ramsey growth framework. In Brito's model the population lives on a straight line. Production and capital accumulation are distributed in space and capital differentials drive the spatial capital dynamics. Boucekkine et al. (2009) further improved and studied this model in the linear utility case. More recent contributions in the same stream are the study of the endogenous growth case by Brito (2012), the characterization of the optimal dynamics of the *AK* model on the a circle *à la Salop* described by Boucekkine et al. (2013)<sup>2</sup>, and the generalization proposed by Aldashev et al. (2014). A different approach considers spatial spillover and excludes capital mobility. This allows a technical simplification of the problem since the diffusion term disappears from the state equation; this is the method chosen for example by Brock and Xepapadeas (2009) and Mossay (2013). A more comprehensive review is available in the introduction of the paper by Brock et al. (2014). Interesting surveys on

---

EPEE, Université d'Evry-Val-d'Essonne (TEPP, FR-CNRS 3126), Département d'Economie, 4 bd. François Mitterrand, 91025 Evry cedex, France, [giorgio.fabbri@univ-evry.fr](mailto:giorgio.fabbri@univ-evry.fr).

I acknowledge fruitful discussions with Raouf Boucekkine (Aix-Marseille School of Economics), Carmen Camacho (CNRS, Paris School of Economics), Giuseppe Freni and Vincenzo Lombardo (University Parthenope, Naples) and Carlo Mantegazza (Scuola Normale Superiore, Pisa). Moreover I thank Tullio Jappelli (CSEF, Naples) and the Committee of the "Unicredit and University research grants" for their useful comments. I also thank the anonymous referee and the associate editor for their thoughtful comments and for providing suggestions that improved the exposition of the paper.

This research was partially supported by the *Post Doc Research Grant of Unicredit & Universities*. It has been developed in the framework of the center of excellence LABEX MME-DII (ANR-11-LABX-0023-01).

<sup>1</sup>Quah (2002) can be considered a precursor of this stream of literature, but in his work the unique production factor ("knowledge") is exogenously generated and no spatial factor mobility is considered.

<sup>2</sup>Brock and Xepapadeas (2008) present a methodological contribution for a class of models including the spatial growth models, focusing again on the circle case.

the subject that contextualize the discussed papers in the wider regional growth literature were proposed by Desmet and Rossi-Hansberg (2010) and Breinlich et al. (2014).

The contribution of the present note is to incorporate the specific role of the geography structure in the growth process and in the *agglomeration vs convergence* long run behavior of the system. The literature refers to several spatial models: straight line, segment and the circle. However, aside from a reflection on the role of the “right” boundary conditions in the state equation, the specific role of the chosen geographic structure has never been considered. As we argue in this paper, there is very likely a technical reason that could explain why this aspect of the problem has not yet been considered.

We present a spatial growth model with the same  $AK$  production function as considered by Boucekkine et al. (2013), with the same law of motion of capital<sup>3</sup>, but with a generic geographic structure. The morphology interacts with the spatial dynamics of the capital and is one determinant of the qualitative optimal behavior of the system. Keeping all other parameters fixed and changing only the geographic structure may lead to a completely different qualitative behavior of the economy. Above all, the convergence found by Boucekkine et al. (2013) is proved to be a particular case of a more complex picture that includes, when varying preferences, parameters and geographical environment, the possibility of long run convergence on the one hand and clustering and agglomeration on the other.

As discussed above, in order to see the whole picture there is a price to pay in terms of mathematical complexity. Indeed, when the general geographic structure is considered, the planner optimization problem leads to an optimal control problem driven by a partial differential equation on a Riemannian manifold. So, in addition to the difficulties of the infinite dimensional structure of the problems that appear in the previously mentioned spatial growth models, the role played by the metric structure of the manifold remains a specific challenge.

After redrafting the model in the form of an optimal control problem (Section 2), we find (Section 3.1) its explicit solution in closed form, describing the optimal dynamics of the spatial distribution of the capital as the solution of a parabolic equation on the geography  $M$ , connected, compact and without boundary. Such a spatio-temporal equation describes the evolution of the economy in the transition towards convergence or agglomeration. The proofs, contained in Appendix A, make use of dynamic programming in an infinite dimensional Hilbert space. This has a specific methodological interest in itself since it is, in our knowledge, the first optimal control problem driven by a diffusion equation on an abstract manifold to be solved and used in the literature.

Our main contribution is contained in Section 3.2, where the role of the geography is developed. In Theorem 3.5 a sharp condition involving the total factor productivity, discount rate, elasticity of inter-temporal substitution, and an index that incorporate the geometric characteristics of the geography, distinguishes situations where the spatial distribution of the (detrended) capital tends to an homogeneous distribution in the long run from situations of long run capital agglomeration and cluster formation.

Section 2 introduces the model setup and presents its main features. Section 3 presents the analytical results and discusses their economic implications. Section 4 includes our overall conclusions and Appendix A contains all the proofs.

---

<sup>3</sup>Initially introduced by Brito (2004).

## 2. THE MODEL

We consider an economy developing on a geography,  $M$ , modeled as an  $n$ -dimensional compact, connected, oriented Riemannian manifold with metric  $g$  and without boundary. Examples of geography with this structure are, for example, the Salop circle or a sphere surface. See Remark 3.6 for comments about our choice of modeling space.

Through the paper we denote a generic spatial point in the geography by  $x \in M$  and the spatial density of capital at point  $x$  and time  $t$  by  $k(t, x)$ . The population is assumed to be constant in time and uniformly distributed, so  $k(t, x)$  is the per-capita distribution of capital. We denote the initial capital distribution by  $k_0(\cdot)$  where  $k_0: M \rightarrow \mathbb{R}$  is then a function representing the initial spatial density of capital for each space point  $x$ .

We consider an  $AK$  production structure for a spatially and temporally homogeneous level of technology  $A$ . If  $\tau(t, x)$  is the trade balance at point  $x$  and time  $t$ , the evolution of the capital is

$$(1) \quad \frac{\partial k(t, x)}{\partial t} = Ak(t, x) - c(t, x) - \tau(t, x)$$

where  $c(t, x) \geq 0$  denotes the consumption at point  $x$  and time  $t$ . Using the same argument as<sup>4</sup> Brito (2004) and several for the one-dimensional spatial case, given a region  $B$  (a connected open subset of  $M$  with regular boundary), the trade balance over  $B$  is given by the sum (i.e., the integral) of what enters each point of the boundary  $\partial B$ :

$$(2) \quad \int_B \tau(t, x) dx = - \int_{\partial B} \frac{\partial k(t, x)}{\partial \hat{n}} dx.$$

When we apply the divergence theorem to (2), we obtain

$$\int_{\partial B} \frac{\partial k(t, x)}{\partial \hat{n}} dx = \int_B \nabla \cdot \nabla k(t, x) dx = \int_B \Delta_x k(t, x) dx,$$

so that, for almost every  $x \in M$ ,

$$(3) \quad -\tau(t, x) = \Delta_x k(t, x),$$

where  $\Delta_x$  is the Laplace-Beltrami operator (we will simply call this *Laplacian* here) on the geography  $M$ . This reduces to the spatial second derivative when we use a one-dimensional model for the space, as is the case for Brito (2004, 2012); Brock and Xepapadeas (2008); Boucekkine et al. (2009, 2013); Aldashev et al. (2014).

Substituting (3) into (1), considering an initial distribution of capital  $k_0(\cdot)$  on  $M$ , and given a consumption profile  $c(\cdot, \cdot)$ , we finally have a partial differential equation on  $M$  that describes the evolution of the capital density  $k(t, x)$ ,

$$(4) \quad \begin{cases} \frac{\partial k(t, x)}{\partial t} = \Delta_x k(t, x) + Ak(t, x) - c(t, x) \\ k(0, x) = k_0(x). \end{cases}$$

The policy maker chooses the consumption  $c(\cdot, \cdot)$  to maximize the following CRRA-Benthamite utility functional

$$(5) \quad J(c(\cdot, \cdot)) = \int_0^{+\infty} e^{-\rho t} \int_M \frac{(c(t, x))^{1-\sigma}}{1-\sigma} dx dt.$$

---

<sup>4</sup>Brito (2004) was the first to adapt an idea arising from classical spatial economics (see, for example, Chapter 8 of Beckmann and Puu, 1985; or Isard et al., 1979) to the benchmark growth models context. The same idea was later used by several other authors, among them Camacho et al. (2008); Brock and Xepapadeas (2008); Boucekkine et al. (2009); Brito (2012); Boucekkine et al. (2013) and Aldashev et al. (2014).

### 3. SPATIO-TEMPORAL DYNAMICS AND CONVERGENCE: THE SOLUTION OF THE MODEL

**3.1. The explicit solution of the model.** We present the solution of the model described above. The proofs are in Appendix A.

We describe the behavior of the aggregate capital,  $K(t) := \int_M k(t, x) dx$ , and of the aggregate consumption,  $C(t) := \int_M c(t, x) dx$ . Their evolution can be sketched by a simple one-dimensional differential equation, as shown by the following proposition.

**Proposition 3.1** *The dynamics of  $K(t)$  is described by*

$$(6) \quad \dot{K}(t) = AK(t) - C(t), \quad K(0) = \int_M k_0(x) dx.$$

Thus, at the aggregate level the model is equivalent to the standard one-dimensional  $AK$  model. We are looking at the internal spatial dynamics of a benchmark  $AK$  model without altering the global structure.

Since we solve the problem using dynamic programming in the Hilbert space  $L^2(M)$  of the square integrable functions<sup>5</sup> from  $M$  to  $\mathbb{R}$  (see Appendix A for details), we make use of the *value function*. For a given initial capital distribution  $k_0(\cdot)$ , we define the value function of our problem starting from  $k_0(\cdot)$  as

$$(8) \quad V(k_0) := \sup_{c(\cdot, \cdot)} J(k_0, c(\cdot, \cdot)),$$

where the supremum is calculated by varying the positive spatio-temporal consumption distributions that ensure the aggregate capital to remain non-negative at any time.  $V(k_0)$  corresponds to the maximal (utilitarian) aggregate welfare that can be guaranteed by the planner for a given initial capital distribution  $k_0(\cdot)$ . Its form can be described explicitly by the following proposition.

**Proposition 3.2** *Suppose that*

$$(9) \quad \rho > A(1 - \sigma),$$

*and consider an initial positive capital distribution  $k_0 \in L^2(M)$ . Then the explicit expression of the value function of the problem at  $k_0$  is*

$$V(k_0) = \frac{1}{1 - \sigma} \left( \frac{\rho - A(1 - \sigma)}{\sigma \text{vol}(M)} \right)^{-\sigma} \left( \int_M k_0(x) dx \right)^{1 - \sigma},$$

where  $\text{vol}(M) := \int_M 1 dx$  is the volume of the geography  $M$ .

We observe that, as in the standard one-dimensional  $AK$  model, the condition (9) is required to ensure the finiteness of the functional and of the value function. Since we chose the dynamic programming approach, we use the characterization of the value function to find the optimal dynamics and the optimal control of the problem. The former is described in the following theorem.

**Theorem 3.3** *Under the hypotheses of Proposition 3.2, the optimal evolution of the capital distribution starting from  $k_0$  is the solution of the following partial differential equation:*

$$(10) \quad \begin{cases} \frac{\partial k(t, x)}{\partial t} = \Delta_x k(t, x) + Ak(t, x) - \left( \frac{\rho - A(1 - \sigma)}{\sigma \text{vol}(M)} \right) \int_M k(t, x) dx \\ k(0, x) = k_0(x). \end{cases}$$

<sup>5</sup>More formally  $L^2(M)$  is the set:

$$(7) \quad L^2(M) := \left\{ f: M \rightarrow \mathbb{R} : \int_M |f(x)|^2 dx < \infty \right\}.$$

It can be endowed with a Hilbert space structure as described in Appendix A.

Equation (10) is a parabolic equation on the geography  $M$  and it describes the optimal evolution of the system from time 0 to  $+\infty$ . The corresponding optimal spatio-temporal consumption can be expressed explicitly as shown in the following proposition.

**Proposition 3.4** *Assume that the hypotheses of Proposition 3.2 are satisfied. Then the optimal consumption is constant in space and exponential in time:  $c(t, x) = c_0 e^{\beta t}$  where  $\beta := \left[ \frac{A-\rho}{\sigma} \right]$  and  $c_0 = \left( \frac{\rho-A(1-\sigma)}{\sigma \text{vol}(M)} \right) K(0)$ , where  $K(0) = \int_M k_0(x) dx > 0$  is the initial level of aggregate capital.*

*Furthermore, the aggregate variables do not have transitional dynamics along the optimal path,  $K(t) = K(0)e^{\beta t}$  and  $C(t) = \left( \frac{\rho-A(1-\sigma)}{\sigma} \right) K(0)e^{\beta t}$ .*

The optimal dynamics of the consumption  $c(t, x)$  is elementary. On the one hand the planner maximizes the utility if, at each time, all individuals in the economy can access the same level of consumption and the per-capita consumption grows exponentially in time. On the other hand, the capital distribution  $k(t, x)$  has a much more elaborate behavior. It is described by the parabolic equation (10) that contains a second order term  $\Delta_x k(t, x)$  and manifests complex transitional dynamics. This dual behavior is common to other infinite dimensional  $AK$  models with CRRA utility<sup>6</sup>. Unlike in other cases, the system can persist in a spatial-unequal capital distribution state. This is discussed further in the subsequent section.

We discussed above that at the aggregate level the model is equivalent to the standard one-dimensional  $AK$  model. Indeed, the previous proposition establishes that the optimal aggregate capital and consumption growth rate are the same as in the one-dimensional case:  $\frac{A-\rho}{\sigma}$ . Moreover the same proportion of the aggregate production is consumed at each time. This global behavior is not reproduced at each spatial point, since, for any fixed  $t \geq 0$ , the solution  $k(t, x)$  of (10) is not a constant function of  $x$ , a part from very specific cases.

**3.2. Geography and convergence.** We study the role of the geographical structure in shaping the long run behavior of the system.

First we need to recall some details about the Laplacian operator  $\Delta_x$  on the geography  $M$ . More details, useful for the proofs, are given in the Appendix A. A non identically zero regular function  $\phi: M \rightarrow \mathbb{R}$  is called an eigenfunction of  $\Delta_x$  if there exists a real number (eigenvalue)  $\lambda$  such that<sup>7</sup>  $\Delta_x \phi = -\lambda \phi$ . It has been proved (see, for example, Chow et al. (2006) page 468) that the set of possible eigenvalues is discrete, and that they form a sequence

$$(11) \quad 0 = \lambda_0 < \lambda_1 < \lambda_2 < \dots < \lambda_k < \dots$$

with  $\lambda_k \xrightarrow{k \rightarrow \infty} +\infty$  and that the constant functions are the unique eigenfunctions associated to the eigenvalue  $\lambda_0 = 0$ .

These values, and in particular the first non-zero eigenvalue of the Laplacian, are important to determine the long run behavior of the spatial growth model. Considering the spatial distribution of the capital in the long run, we discount it by the growth rate of the aggregate variables, thus obtaining the *detrended* spatial distribution of the capital at time  $t$   $k_D(t, x)$

$$k_D(t, x) := e^{-\beta t} k(t, x).$$

If we divide  $k_D$  by  $K(0)$ , we obtain exactly the density of the portion of aggregate capital localized at point  $x$  at time  $t$ . The following theorem shows explicitly how the technological

<sup>6</sup>See, for example, Boucekkine et al. (2005); Fabbri and Gozzi (2008); Boucekkine et al. (2010, 2013). In the first three of these works the capital accumulation takes the form of a delay differential equation. However, although the mathematical structure is somewhat different, the same dual behavior in the evolution of capital and consumption is reproduced.

<sup>7</sup>We define here an eigenvalue as a number  $\lambda$  such that  $\Delta_x \phi = -\lambda \phi$ . This is the standard in the differential geometry literature where researchers often study the operator  $(-\Delta_x)$  rather than  $\Delta_x$ .

and preferences parameters and the geographic characteristics interact to determine the spatial convergence or capital cluster formation.

**Theorem 3.5** *Assume that the hypotheses of Proposition 3.2 hold. Then<sup>8</sup> the detrended capital distribution  $k_D(t, x)$  tends to a spatially equally distributed state if and only if*

$$(12) \quad \rho < A(1 - \sigma) + \sigma\lambda_1,$$

*otherwise we obtain a long-run spatial capital agglomeration.*

Theorem 3.5 is the key outcome of this paper. In Theorem A.10 and Remark A.11 in Appendix A, we present a more detailed result, showing the long run level of the detrended capital in the case of convergence and characterizing for agglomeration across locations, the limit of the detrended capital in terms of (non-constant) eigenfunctions of the Laplacian. However, in (12) we already have all the important ingredients to see how the various elements of the models interact to determine the convergence or the agglomeration in the long run behavior of the system.

A greater consumption impatience, measured by the discount rate  $\rho$ , tends to prevent convergence: increasing  $\rho$  pushes the consumption level  $c_0$ , characterized in Proposition 3.4. It is the same at each spatial point but it is relatively higher in the depressed areas, being more affected by the fall in investment levels.

The role of  $A$  depends on its impact on the consumption level of  $c_0$  and it changes depending on the value of  $\sigma$ , the inverse of the elasticity of intertemporal substitution. This is not surprising: in the one-dimensional  $AK$  model, the effect of  $A$  on the level of the consumption in terms of physical capital,  $\left(\frac{\rho - A(1 - \sigma)}{\sigma}\right)$ , varies if  $\sigma$  is greater or lower than 1. Realistic values of the elasticity of intertemporal substitution are well below the unity, so the corresponding values of  $\sigma$  are greater than 1. Consequently the term  $A(1 - \sigma)$  is negative and an increase of the total factor productivity level diminishes the possibility of convergence. Thus, the prevailing effect is a differential push: the impact of a gain in the total factor productivity on the production at a point  $x$  is proportional to the capital at that point and stronger in richer areas.

$\lambda_1$  is the first non-zero eigenvalue of the Laplacian on the geography  $M$  and summarizes its geometric properties. It can be shown that, for geographies with the same volume,  $\lambda_1$  is smaller if  $M$  presents narrower bottlenecks (see Buser, 1982, Theorem 1.2 for a detailed result). Hence,  $\lambda_1$  decreases as the ecological barriers and geographical obstacles to capital diffusion increase and, consistently, it is more difficult to verify (12).

Increasing the intertemporal elasticity of substitution  $1/\sigma$ , in the case of a positive growth rate (i.e.  $A > \rho$ ), increases the part of production used for investment for each point in space and time. The consequent effect depends on which of the two effects is stronger. The divergent effect is measured by  $A$  and is due to different gains in production given different capital densities or the homogenizing effect of capital spread quantified by  $\lambda_1$ .

**Some numerical examples.** To underline the role of the geographical structure in Theorem 3.5, we consider some numerical examples. They are summarized in Table 1 and they concern two possible geographies of dimension  $n = 2$  (surfaces). The first (Example 1) is a sphere of radius 1. Its (2-dimensional) volume is  $4\pi$  and the value of  $\lambda_1$  is 2 (see, for example, Theorem 22.1, page 169 of Shubin, 2013). The second (Examples 2 and 3) is a surface-preserving deformation of the same sphere with an apparent bottleneck and  $\lambda_1 = 0.1$  (tightening the bottleneck we can obtain geographies with smaller  $\lambda_1$ ).

<sup>8</sup>Apart from the particular case where the initial capital distribution  $k_0$  is constant in space. In this case, even if condition (12) is violated, the detrended capital distribution remains always constant.

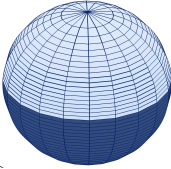
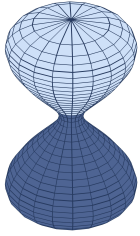
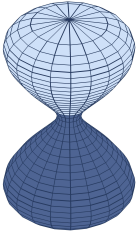
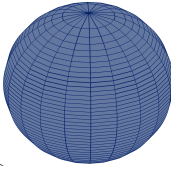
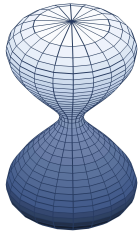
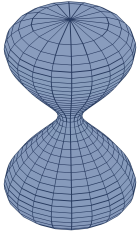
	<b>Example 1</b>	<b>Example 2</b>	<b>Example 3</b>
	$A = 0.14,$ $\sigma = 3,$ $\rho = 5\%$	$A = 0.14,$ $\sigma = 3,$ $\rho = 5\%$	$A = 0.14,$ $\sigma = 3,$ $\rho = 2\%$
	$\lambda_1 = 2$	$\lambda_1 = 0.1$	$\lambda_1 = 0.1$
Initial detrended capital distribution			
Final detrended capital distribution			
	convergence	agglomeration	convergence

TABLE 1. Convergence and agglomeration effects in some numerical examples (the shades of blue represent the different levels of the detrended capital)

In all examples we fixed the same initial capital distribution, equal to 0.8 in the lower half of the surface and to 0.2 in the upper.

In Examples 1 and 2 we keep the same values for  $A$ ,  $\sigma$ , and  $\rho$  and only change the geography. As a consequence, condition (12) is verified in Example 1 and violated in Example 2. So the long run distribution of the detrended capital is constant in the first case (convergence) and non-constant in the second (agglomeration).

In Example 2 and 3 we keep the same geographical structure, but we change the parameter  $\rho$ , causing condition (12) to be satisfied in Example 3.

**Remark 3.6** As remarked in Section 5 of Breinlich et al. (2014), we can recognize two families of theoretical spatial growth models: those where the space is *ordered*, i.e. the interaction among locations depend on their distance; and those where the space is *non-ordered*. In our model (as, for example, in those of Brito 2004, 2012 or Boucekkine et al. 2009, 2013) the space is ordered. Indeed the capital diffusion process among locations is driven by a heat equation; this implies that exchanges and fluxes among points become more intense as they get closer<sup>9</sup>.

We focus on the case of a connected geography without boundary. However, the structure of the geography can include a series of regions among which, given some obstacles, the capital has some specific impediment to flow. This is the case for Examples 2 and 3 where North and South regions can be clearly identified.

<sup>9</sup>The distance is defined in terms of the metric  $g$  on  $M$ .



In the literature discussed above, to study the spatial dynamics of a certain regions in relation with some exterior economy, sometimes (one dimensional) models with boundary have been considered (see, for example, Brito, 2012 or Boucekkine et al., 2009). Apart from the difficulty to solve the optimization program for a generic boundary condition, the problem is to understand which are the “right” conditions that correspond to our economic intuition: do we need to impose some exogenous conditions on the value of the capital at the boundary (Dirichlet boundary conditions), or on its derivative (Neumann boundary conditions), or none? It is difficult to find a clear argument to prefer one condition to another. In different contributions different approaches are used and often the choice fell on the study of problems with zero-boundary conditions, i.e. where the capital or its spatial normal derivative are assumed to be zero on the boundary.

In our proposed framework, regions can be described by the geographical structure, there are no exogenous boundary conditions among them and the flow of the capital is endogenously characterized.

Our setting imposes a restriction to a bounded geography but one can wonder what happens if we decide to “scale” the geography i.e. increase its volume while keeping its form. Proposition 3.4, in terms of optimal growth rates, remains exactly the same<sup>10</sup>.

The effect of scaling on the convergence/agglomeration analysis is more interesting. One can show (see Chapter 1 of the book by Chavel, 1984) that, if we rescale a geography  $M$  to get a geography with volume  $r$  times the original, the new eigenvalues are the old ones multiplied by a factor  $r^{-\frac{2}{n}}$  ( $n$  being the dimension of the geography). Hence, if  $r > 1$ , they reduce. In particular, if we take large values of  $r$ , condition (12) can be satisfied only by a restricted choice of parameters. In particular taking the limit  $r \rightarrow \infty$ , condition (12) becomes  $\rho < A(1 - \sigma)$ , which is not compatible with (9).  $\square$

**Example 3.7** As an example, consider the case studied by Boucekkine et al. (2013). The geography is the Salop circle of radius 1, so  $vol(M) = 2\pi$  (the circumference) and  $\lambda_1 = 1$  (see Theorem 22.1, page 169 of Shubin, 2013). Condition (12) becomes

$$(13) \quad \rho < A(1 - \sigma) + \sigma,$$

which is precisely condition (13) of Boucekkine et al. (2013), in particular if we restrict our attention to the values of  $\rho$ ,  $A$  and  $\sigma$  satisfying such an inequality, we reproduce exactly their convergence result (Theorem 3.3 of their paper).

The circle can be described in polar coordinates using the parameter  $\theta \in [0, 2\pi]$  (with  $\theta = 0$  and  $\theta = 2\pi$  identified). We include the initial condition, given by  $k_0(\theta) = 10$  if  $\theta \in [0, \pi)$  and  $k_0(\theta) = 20$  if  $\theta \in [\pi, 2\pi)$ . If we consider a parametrization where  $\rho < A(1 - \sigma) + \sigma$ , then, from Theorem A.10, part 1, the long run homogeneous outcome  $\bar{k}(\theta) = \frac{\int_0^{2\pi} k_0(\alpha) d\alpha}{2\pi} = \frac{10\pi + 20\pi}{2\pi} = 15$ , for any  $\theta \in [0, 2\pi)$ . If the parameters are chosen to guarantee that  $\rho = A(1 - \sigma) + \sigma$ , then we have the second graphical example of Boucekkine et al. (2013). They only study it numerically, whereas we can characterize the solution analytically, finding the non-homogeneous long-run distribution of the detrended capital (from Theorem A.10 part 2)

$$\bar{k}(\theta) = \frac{\int_0^{2\pi} k_0(\alpha) d\alpha}{2\pi} + \frac{1}{\pi} \left( \int_0^{2\pi} k_0(\alpha) \sin(\alpha) d\alpha \right) \sin(\theta) = 15 + \frac{20}{\pi} \sin(\theta), \quad \theta \in [0, 2\pi).$$

Finally, if  $\rho > A(1 - \sigma) + \sigma$  (Theorem A.10 part 3), then only the non-constant addendum remains in the long-run distribution.

<sup>10</sup>The levels of the variables can change, depending on how we decide to rescale the population and the initial endowment. If we rescale the population and leave the same per-capita endowment, then the per-capita levels remain constant.

If we restrict our attention to a circle of radius 1, any reasonable choice of the parameters (for a reasonable growth rate and discounting rate, and  $\sigma \gg 1$ ) verifies the convergence condition. However, changing the geography, the value of  $\lambda_1$  can become as close to zero as we wish (see also Randol, 1974). Thus, the choice of possible parameters satisfying both conditions (9) and (12) can be very small.  $\square$

#### 4. CONCLUSIONS

The contribution of the present work is to investigate the role of geography in the evolution of a spatial growth model. To this extent, we consider an *AK* spatial model with capital mobility and a generic geographic structure.

Our main finding is that changing the geography changes the qualitative behavior of the system. Keeping the same parameters for preferences, discount rate, and total factor productivity, we observe convergence or agglomeration of the detrended capital across the locations depending on the geography structure. We have precisely characterized the analytical conditions that lead to different qualitative behaviors.

#### REFERENCES

- G. Aldashev, S. Aldashev, and T. Carletti. On convergence in the spatial ak growth models. *arXiv preprint arXiv:1401.4887*, 2014.
- M. Beckmann and T. Puu. *Spatial economics: density, potential, and flow*. North-Holland, 1985.
- A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter. *Representation and control of infinite dimensional systems*. Birkhäuser Boston Inc., Boston, MA, 2007.
- R. Boucekkine, O. Licandro, L. A. Puch, and F. del Rio. Vintage capital and the dynamics of the AK model. *Journal of Economic Theory*, 120(1):39–72, 2005.
- R. Boucekkine, C. Camacho, and B. Zou. Bridging the gap between growth theory and the new economic geography: The spatial Ramsey model. *Macroeconomic Dynamics*, 13(1):20–45, 2009.
- R. Boucekkine, G. Fabbri, and F. Gozzi. Maintenance and investment: Complements or substitutes? a reappraisal. *Journal of Economic Dynamics and Control*, 34(12):2420–2439, 2010.
- R. Boucekkine, C. Camacho, and G. Fabbri. Spatial dynamics and convergence: The spatial ak model. *Journal of Economic Theory*, 148(6):2719–2736, 2013.
- H. Breinlich, G. I. Ottaviano, and J. R. Temple. Regional Growth and Regional Decline. In *Handbook of Economic Growth, Vol. 2*, chapter 4, pages 683–779. Elsevier, 2014.
- P. Brito. The Dynamics of Growth and Distribution in a Spatially Heterogeneous World. Working Papers Department of Economics 2004/14, ISEG, University of Lisbon, 2004.
- P. Brito. Global endogenous growth and distributional dynamics. Munich Personal RePEc Archive, Paper No. 41653, October 2012.
- W. Brock and A. Xepapadeas. Diffusion-induced instability and pattern formation in infinite horizon recursive optimal control. *Journal of Economic Dynamics and Control*, 32(9):2745–2787, 2008.
- W. Brock and A. Xepapadeas. General pattern formation in recursive dynamical systems models in economics. Working Papers 2009.49, Fondazione Eni Enrico Mattei, Jun 2009.
- W. Brock, A. Xepapadeas, and A. Yannacopoulos. Optimal Control in Space and Time and the Management of Environmental Resources. DEOS Working Papers 1402, Athens University of Economics and Business, Feb 2014.
- P. Buser. A note on the isoperimetric constant. In *Annales scientifiques de l'École Normale Supérieure*, volume 15, pages 213–230. Société mathématique de France, 1982.
- C. Camacho, B. Zou, and M. Briani. On the dynamics of capital accumulation across space. *European Journal of Operational Research*, 186(2):451–465, 2008.
- I. Chavel. *Eigenvalues in Riemannian geometry*, volume 115 of *Pure and Applied Mathematics*. Academic Press Inc., Orlando, FL, 1984. Including a chapter by Burton Randol, With an appendix by Jozef Dodziuk.
- B. Chow, P. Lu, and L. Ni. *Hamilton's Ricci flow*, volume 77 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2006.
- K. Desmet and E. Rossi-Hansberg. On spatial dynamics. *Journal of Regional Science*, 50(1):43–63, 2010.
- G. Fabbri and F. Gozzi. Solving optimal growth models with vintage capital: The dynamic programming approach. *Journal of Economic Theory*, 143(1):331–373, 2008.
- A. Grigor'yan. *Heat kernel and analysis on manifolds*. American Mathematical Soc., 2012.
- W. Isard, P. Liossatos, Y. Kanemoto, and P. C. Kaniss. *Spatial dynamics and optimal space-time development*. North-Holland, 1979.

- P. Mossay. A theory of rational spatial agglomerations. *Regional Science and Urban Economics*, 43(2): 385–394, 2013.
- D. Quah. Spatial agglomeration dynamics. *The American Economic Review*, 92(2):247–252, 2002.
- B. Randol. Small eigenvalues of the laplace operator on compact riemann surfaces. *Bulletin of the American Mathematical Society*, 80(5):996–1000, 1974.
- M. A. Shubin. *Pseudodifferential operators and spectral theory*. Springer, 2013.
- A. Smith. *An inquiry into the nature and causes of the wealth of nations*. First edition, 1776.

## APPENDIX A. PROOFS

**A.1. Notations and preliminary results.** For the reader convenience, we first provide some definitions and preliminary results necessary to the proofs.

We start by recalling how to rewrite the optimal control problem stated in Section 2 as an optimal control problem in the space of square integrable functions from  $M$  to  $\mathbb{R}$ . The space  $L^2(M)$ , defined in (7), is a Hilbert space, we denote by  $\langle f, g \rangle := \int_M f(x)g(x) dx$  its scalar product. We define the operator  $G$  on  $L^2(M)$  as follows<sup>11</sup>

$$\begin{cases} D(G) := H^2(M) \\ G(f) = \Delta_x f. \end{cases}$$

$G$  is the (self-adjoint) generator of the heat semigroup on  $L^2(M)$ . It is a  $C_0$  semigroup on  $L^2(M)$  (see Section 4.3 of Grigor'yan, 2012 and Bensoussan et al., 2007 for the general theory of  $C^0$ -semigroups). The state equation (4) can be rewritten as an evolution equation in the Hilbert space  $L^2(M)$  as follows:

$$(14) \quad \begin{cases} \dot{k}(t) = Gk(t) + Ak(t) - c(t) \\ k(0) = k_0 \end{cases}$$

where  $k(t)$  and  $c(t)$  are interpreted as the functions of variable  $x$  defined by  $k(t)(x) \equiv k(t, x)$  and  $c(t)(x) \equiv c(t, x)$ . The mild solution of (14), see Definition 3.1, page 129 of Bensoussan et al. (2007), is given by

$$(15) \quad k_{k_0, c}(t) = e^{Gt} k_0 + \int_0^t e^{(t-s)G} (Ak_{k_0, c}(s) - c(s)) ds$$

or, called  $\tilde{G} := G + A$  ( $G$  plus  $A$  times the identity operator),

$$(16) \quad k_{k_0, c}(t) = e^{\tilde{G}t} k_0 - \int_0^t e^{(t-s)\tilde{G}} c(s) ds.$$

Observe that, chosen a control  $c$ , the aggregate capital is given by  $K(t) = \langle k_{k_0, c}(t), \mathbb{1} \rangle$  where  $\mathbb{1} : M \rightarrow \mathbb{R}$  is the function that is identically equal to 1. We use the notation  $L_{loc}^2(0, +\infty; L^2(M))$  for the following functions space

$$L_{loc}^2(0, +\infty; L^2(M)) := \left\{ f : [0, +\infty) \times M \rightarrow \mathbb{R} : \int_0^T \int_M |f(t, x)|^2 dx dt < \infty \forall T > 0 \right\}.$$

The set of admissible controls of the optimal control problem (4)-(5) can be then written as

$$(17) \quad \mathcal{U}_{k_0} := \{c(\cdot, \cdot) \in L_{loc}^2(0, +\infty; L^2(M)) : c(\cdot, \cdot) \geq 0 \text{ and } K(t) > 0 \text{ for } t \geq 0\}.$$

So the optimal control problem described in Section 2 is equivalent to the problem of maximizing the functional

$$J(k_0, c) := \int_0^{+\infty} e^{-\rho t} \langle \mathbb{1}, U(c(t)) \rangle dt,$$

where  $U(\eta)(x) = \frac{(\eta(x))^{1-\sigma}}{1-\sigma}$ , among the controls of (17), subject to (14).

As already mentioned in Section 3, we call eigenfunction of  $\Delta_x$  a (non identically zero) regular function  $\phi : M \rightarrow \mathbb{R}$  such that  $\Delta_x \phi = -\lambda \phi$  for a real number (eigenvalue)  $\lambda$  and it can be proved (see *e.g.* Chow et al. (2006) page 468) that the set of the possible eigenvalues is discrete and they form a sequence  $0 = \lambda_0 < \lambda_1 < \lambda_2 < \dots < \lambda_k < \dots$ . We call eigenspace associated with the eigenvalue  $\lambda_n$  the vector space of the eigenfunctions associated with the eigenvalue  $\lambda_n$  and we denote it by  $S_n$ . It can be proved that the dimension of  $S_n$  is finite (see *e.g.* Chow et al., 2006, page 469), we denote it by  $\theta_n$ . It is possible to choose an orthonormal basis of  $L^2(M)$  of (normalized) eigenfunctions of the Laplacian  $\phi_n^j$ , for  $n \in \mathbb{N}$  and  $j \in \{1, \dots, \theta_n\}$ , where, for any  $n \geq 0$ ,  $\phi_n^1, \dots, \phi_n^{\theta_n}$  are eigenfunctions associated to the eigenvalue  $\lambda_n$ . Any  $f \in L^2(M)$  can be written as the  $L^2(M)$ -limit of the series  $f = \sum_{n=0}^{\infty} \sum_{j=1}^{\theta_n} \langle f, \phi_n^j \rangle \phi_n^j$  and  $\|f\|_{L^2(M)} = \sum_{n=0}^{\infty} \sum_{j=1}^{\theta_n} \langle f, \phi_n^j \rangle^2$ .

<sup>11</sup>The Sobolev space  $H^2(M)$  is the completion, w.r.t. the norm  $\|f\|_{H^2} := \left( \sum_{i=0}^2 \int_M |\nabla^i f(x)|^2 dx \right)^{1/2}$  of the space of the  $C^\infty$  functions (see Grigor'yan, 2012, Section 4 for details). We write  $D(G)$  to denote the *domain* of the operator  $G$ .

It can also be shown that the dimension of  $S_0$  is exactly 1 and it contains only constant functions *i.e.* the functions of the form  $\alpha \mathbb{1}$  for some  $\alpha \in \mathbb{R}$ . In particular  $e^{Gt} \mathbb{1} = \mathbb{1}$  and  $e^{\tilde{G}t} \mathbb{1} = e^{-At} \mathbb{1}$ . The unique normalized function (w.r.t.  $L^2$ -norm) of  $S_0$  is  $\phi_0 = \frac{1}{\sqrt{\text{vol}(M)}} \mathbb{1}$ .

*Proof of Proposition 3.1.* Given an initial datum  $k_0$  and chosen an admissible control  $c(\cdot, \cdot)$ , using (16), we have

$$K(t) = \langle k_{k_0, c}(t), \mathbb{1} \rangle = \left\langle e^{\tilde{G}t} k_0, \mathbb{1} \right\rangle - \int_0^t \left\langle e^{(t-s)\tilde{G}} c(s), \mathbb{1} \right\rangle ds.$$

Since  $G$  and then  $e^{\tilde{G}t}$  are self-adjoint the expression above equals

$$\left\langle k_0, e^{\tilde{G}t} \mathbb{1} \right\rangle - \int_0^t \left\langle c(s), e^{(t-s)\tilde{G}} \mathbb{1} \right\rangle ds$$

and, since  $e^{\tilde{G}t} \mathbb{1} = e^{-tA} \mathbb{1}$  and the aggregate consumption is given by  $C(t) = \langle c(t), \mathbb{1} \rangle$ , we obtain

$$K(t) = e^{-tA} K(0) - \int_0^t e^{-A(t-s)} C(s) ds$$

so  $K(t)$  is exactly the solution of (6). This proves the claim.  $\square$

**A.2. Proofs of results of Section 3.1.** In this subsection we solve the optimal control problem using dynamic programming in infinite dimensions and we prove Proposition 3.2, Theorem 3.3 and Proposition 3.4.

First, we observe in the following proposition that the condition (9) demanded in Proposition 3.2 (and then in all the subsequent results) is sufficient to ensure the finiteness of the value function. In fact what we prove here is just that  $V < +\infty$ . The other bound will be a corollary of the following results since the utility along the optimal trajectory will be bigger than  $-\infty$ , and so will the supremum of the utility varying the control.

**Proposition A.1** *If*

$$(18) \quad \rho > A(1 - \sigma)$$

*then all the trajectories give a bounded utility from above, more precisely, for all positive  $k_0 \in L^2(M)$ ,  $V(k_0) := \sup_{c \in \mathcal{U}_{k_0}} J(c(\cdot, \cdot)) < +\infty$ .*

*Proof of Proposition A.1.* Since the claim is obvious if  $\sigma > 1$  we prove it only for  $\sigma \in (0, 1)$ . Observe first that, from (6), since  $C(\cdot) \geq 0$ , we have

$$(19) \quad K(t) \leq K(0)e^{At}.$$

Using Jensen inequality on the space  $[0, +\infty) \times M$  with the measure  $\frac{1}{\rho \text{vol}(M)} e^{-\rho t} dt \otimes dx$ , (6) and then integration by part, we have

$$(20) \quad \begin{aligned} \frac{1}{\rho \text{vol}(M)} \int_0^{+\infty} \int_M e^{-\rho t} c(t, x)^{1-\sigma} dx dt &\leq \left( \frac{1}{\rho \text{vol}(M)} \int_0^{+\infty} e^{-\frac{\rho}{1-\sigma} t} C(t) dt \right)^{1-\sigma} \\ &= \left( \frac{1}{\rho \text{vol}(M)} \int_0^{+\infty} e^{-\frac{\rho}{1-\sigma} t} (AK(t) - \dot{K}(t)) dt \right)^{1-\sigma} \\ &\leq \left( \frac{1}{\rho \text{vol}(M)} \int_0^{+\infty} e^{-\frac{\rho}{1-\sigma} t} AK(t) dt - \left| K(t) e^{-\frac{\rho}{1-\sigma} t} \right|_{t=0}^{t=+\infty} - \int_0^{+\infty} \frac{\rho}{1-\sigma} e^{-\frac{\rho}{1-\sigma} t} K(t) dt \right)^{1-\sigma}. \end{aligned}$$

Applying (18) and (19) the last expression can be easily seen to be lower than a constant independent of the control  $c$ . So we have the claim.  $\square$

We now study the optimal control problem using the dynamic programming approach in the space  $L^2(M)$ . So first we write, in (21), the Hamilton-Jacobi-Bellman (HJB) equation of the problem then we find an explicit solution (Proposition A.3) and we use such a solution to derive, in (23), a feedback. Eventually we demonstrate (Proposition A.8) that such a feedback is optimal and that the solution of the HJB equation we found is in fact the value function of the problem. These two last results prove Proposition 3.2 and Theorem 3.3. The proof of Proposition 3.4 follows as a corollary.

The HJB equation associated to our problem is<sup>12</sup>:

$$(21) \quad \rho v(k) = \langle k, G\mathcal{D}v(k) \rangle + A \langle k, \mathcal{D}v(k) \rangle + \sup_{c \geq 0} \{ - \langle c, \mathcal{D}v(k) \rangle + \langle \mathbb{1}, U(c) \rangle \}$$

where  $\mathcal{D}v$  represents the Fréchet differential of the function  $v: L^2(M) \rightarrow \mathbb{R}$ .

<sup>12</sup>Observe that we read the term  $\langle Gk, \mathcal{D}v(k) \rangle$  appearing in the usual expression of the HJB equation as  $\langle k, G^* \mathcal{D}v(k) \rangle$ . Since  $G$  is self-adjoint, it equals  $\langle k, G\mathcal{D}v(k) \rangle$  and we find the form given in (21).

**Definition A.2** Let  $O \subseteq L^2(M)$  an open set.  $v: O \rightarrow \mathbb{R}$  is a solution of (21) on  $O$  if  $v \in C^1(O)$ ,  $\mathcal{D}v \in C(O, D(G))$  and  $v$  solves pointwise (21) on  $O$ .

**Proposition A.3** The function

$$(22) \quad v(k) = \alpha \langle k, \mathbb{1} \rangle^{1-\sigma},$$

with  $\alpha = \frac{1}{1-\sigma} \left( \frac{\rho - A(1-\sigma)}{\sigma \text{vol}(M)} \right)^{-\sigma}$ , is a solution of (21) on the halfspace  $\Omega := \{f \in L^2(M) : \langle \mathbb{1}, f \rangle > 0\}$ .

*Proof.* We verify the statement directly. We observe that  $\mathcal{D}v(k) = \alpha(1-\sigma) \langle k, \mathbb{1} \rangle^{-\sigma} \mathbb{1}$  so the candidate-solution is a solution of (21) if and only if

$$\rho \alpha \langle k, \mathbb{1} \rangle^{1-\sigma} = \alpha(1-\sigma) \langle k, \mathbb{1} \rangle^{-\sigma} \langle k, G\mathbb{1} \rangle + A\alpha(1-\sigma) \langle k, \mathbb{1} \rangle^{-\sigma} \langle k, \mathbb{1} \rangle + \sup_{c \geq 0} \left\{ -\alpha(1-\sigma) \langle k, \mathbb{1} \rangle^{-\sigma} \langle c, \mathbb{1} \rangle + \langle \mathbb{1}, U(c) \rangle \right\}.$$

Observing that  $G\mathbb{1} = 0$  and that the supremum is attained when  $c = (\alpha(1-\sigma))^{-1/\sigma} \langle k, \mathbb{1} \rangle \mathbb{1}$  the expression above becomes:

$$\rho \alpha \langle k, \mathbb{1} \rangle^{1-\sigma} = A\alpha(1-\sigma) \langle k, \mathbb{1} \rangle^{1-\sigma} - \text{vol}(M)\alpha(1-\sigma) (\alpha(1-\sigma))^{-1/\sigma} \langle k, \mathbb{1} \rangle^{1-\sigma} + \text{vol}(M) \frac{\left[ (\alpha(1-\sigma))^{-1/\sigma} \langle k, \mathbb{1} \rangle \right]^{1-\sigma}}{1-\sigma}$$

simplifying the non-zero factor  $\alpha \langle k, \mathbb{1} \rangle^{1-\sigma}$  the previous expression is equivalent to

$$\rho = A(1-\sigma) - \text{vol}(M)(1-\sigma) (\alpha(1-\sigma))^{-1/\sigma} + \text{vol}(M) (\alpha(1-\sigma))^{-1/\sigma}.$$

Using the explicit expression of  $\alpha$  given in the statement, we can easily see that the previous equation is verified. This proves the claim.  $\square$

**Definition A.4** Given  $O$  an open subset of  $L^2(M)$ , a function  $\Psi: O \rightarrow L^2(M)$  is said to be a feedback in  $O$  if, for any  $k_0 \in O$ , the equation

$$\begin{cases} \dot{k}(t) = Gk(t) + Ak(t) - \Psi(k(t)) \\ k(0) = k_0 \end{cases}$$

has a unique solution  $k_{\Psi, k_0}(\cdot)$  and  $k_{\Psi, k_0}(t) \in O$  for all  $t \geq 0$ .

**Definition A.5** Given  $O$  an open subset of  $L^2(M)$ , a function  $\Psi: O \rightarrow L^2(M)$  is said to be an admissible feedback in  $O$  if it is a feedback in  $O$  and, for any  $k_0 \in O$ ,  $\Psi(k_{\Psi, k_0}(\cdot)) \in \mathcal{U}_{k_0}$ .

**Definition A.6** Given  $O$  an open subset of  $L^2(M)$  a function  $\Psi: O \rightarrow L^2(M)$  is said to be an optimal feedback in  $O$  if it is an admissible feedback in  $O$  and, for any  $k_0 \in O$ ,  $\Psi(k_{\Psi, k_0}(\cdot))$  is an optimal control.

The feedback associated to the solution of the HJB equation found in Proposition A.3 is given by

$$(23) \quad \begin{cases} \Phi: \Omega \rightarrow L^2(M) \\ \Phi: k \mapsto \sup_{c \in L^2(M), c \geq 0} \left\{ -\alpha(1-\sigma) \langle k, \mathbb{1} \rangle^{-\sigma} \langle c, \mathbb{1} \rangle + \langle \mathbb{1}, U(c) \rangle \right\} = \left( \frac{\rho - A(1-\sigma)}{\sigma \text{vol}(M)} \right) \langle k, \mathbb{1} \rangle \mathbb{1}. \end{cases}$$

**Proposition A.7**  $\Phi$ , defined in (23), is an admissible feedback in  $\Omega$  (defined in Proposition A.3). More precisely, for any  $k_0 \in \Omega$ , if we define  $K_{\Phi, k_0}(t) := \langle \mathbb{1}, k_{\Phi, k_0}(t) \rangle$  we have

$$(24) \quad K_{\Phi, k_0}(t) = K(0)e^{\beta t}$$

where  $K(0) = \langle \mathbb{1}, k_0 \rangle$  and

$$\beta := \left[ \frac{A - \rho}{\sigma} \right].$$

*Proof.* We choose  $k_0 \in \Omega$  and to lighten the notation, we write  $k_{\Phi}$  instead of  $k_{\Phi, k_0}$  and  $K_{\Phi}$  instead of  $K_{\Phi, k_0}$ . Using (15) and replacing  $c(t)$  by the feedback  $\Phi(k_{\Phi}(t))$  we have

$$k_{\Phi}(t) = e^{Gt}k_0 + \int_0^t e^{(t-s)G} (Ak_{\Phi}(s) - c(s)) ds = e^{Gt}k_0 + \int_0^t e^{(t-s)G} \left( Ak_{\Phi}(s) - \frac{\rho - A(1-\sigma)}{\sigma \text{vol}(M)} \langle k_{\Phi}(s), \mathbb{1} \rangle \mathbb{1} \right) ds,$$

so

$$K_{\Phi}(t) = \left\langle k_0, e^{Gt} \mathbb{1} \right\rangle + \int_0^t \left\langle \left( Ak_{\Phi}(s) - \frac{\rho - A(1-\sigma)}{\sigma \text{vol}(M)} \langle k_{\Phi}(s), \mathbb{1} \rangle \mathbb{1} \right), e^{(t-s)G} \mathbb{1} \right\rangle ds.$$

Given that  $e^{Gt} \mathbb{1} = \mathbb{1}$ , the expression above becomes:

$$(25) \quad \begin{aligned} K_{\Phi}(t) &= \langle k_0, \mathbb{1} \rangle + \int_0^t \left\langle \left( Ak_{\Phi}(s) - \frac{\rho - A(1-\sigma)}{\sigma \text{vol}(M)} \langle k_{\Phi}(s), \mathbb{1} \rangle \mathbb{1} \right), \mathbb{1} \right\rangle ds \\ &= K_{\Phi}(0) + \int_0^t K_{\Phi}(s) \left[ A - \text{vol}(M) \frac{\rho - A(1-\sigma)}{\sigma \text{vol}(M)} \right] ds = K_{\Phi}(0) + \int_0^t K_{\Phi}(s) \left[ \frac{A - \rho}{\sigma} \right] ds \end{aligned}$$

and the claim is proved.  $\square$

**Proposition A.8** *Assume that (18) is satisfied. Then  $\Phi$  defined in (23) is an optimal feedback in  $\Omega$  and the value function of the problem computed at  $k_0$  is*

$$V(k_0) = \frac{1}{1-\sigma} \left( \frac{\rho - A(1-\sigma)}{\sigma \operatorname{vol}(M)} \right)^{-\sigma} \langle k_0, \mathbb{1} \rangle^{1-\sigma}$$

*Proof.* Call  $c^*(t) := \Phi(k_\Phi(t))$ . To prove that  $c^*(\cdot)$  is an optimal control, we have to prove that for any other admissible control  $\tilde{c}(\cdot)$  ( $\tilde{k}(\cdot)$  being the related trajectory),  $J(k_0, c^*) \geq J(k_0, \tilde{c})$ . Denote by  $w(t, k): \mathbb{R} \times L^2(M) \rightarrow \mathbb{R}$  the function  $w(t, k) := e^{-\rho t} v(k)$ . If we fix  $T > 0$ , we have:

$$\begin{aligned} v(k_0) - w(T, \tilde{k}(T)) &= w(t, \tilde{k}(0)) - w(T, \tilde{k}(T)) = - \int_0^T \frac{d}{dt} w(t, \tilde{k}(t)) dt \\ &= \int_0^T e^{-\rho t} \left[ \rho v(\tilde{k}(t)) - \left\langle G\tilde{k}(t) + A\tilde{k}(t) - \tilde{c}(t), \mathcal{D}v(\tilde{k}(t)) \right\rangle \right] dt. \end{aligned}$$

The last expression makes sense thanks to the regularizing properties of the heat semigroup: for any  $t > 0$ ,  $\tilde{k}(t) \in D(G)$ . Using (19), the explicit form of  $v$  given in (22) and the hypothesis (12) we can easily see that  $w(T, \tilde{k}(T)) \rightarrow 0$  when  $T \rightarrow \infty$  so we can pass to the limit in the previous equation and find that

$$v(k_0) = \int_0^{+\infty} e^{-\rho t} \left[ \rho v(\tilde{k}(t)) - \left\langle A\tilde{k}(t) - \tilde{c}(t), \mathcal{D}v(\tilde{k}(t)) \right\rangle - \left\langle \tilde{k}(t), G\mathcal{D}v(\tilde{k}(t)) \right\rangle \right] dt$$

and then

$$\begin{aligned} v(k_0) - J(k_0, \tilde{c}) &= \int_0^{+\infty} e^{-\rho t} \left[ \left( \rho v(\tilde{k}(t)) - \left\langle A\tilde{k}(t), \mathcal{D}v(\tilde{k}(t)) \right\rangle - \left\langle \tilde{k}(t), G\mathcal{D}v(\tilde{k}(t)) \right\rangle \right) \right. \\ &\quad \left. + \left( \left\langle \tilde{c}(t), \mathcal{D}v(\tilde{k}(t)) \right\rangle - \langle \mathbb{1}, U(\tilde{c}(t)) \rangle \right) \right] dt \\ &= \int_0^{+\infty} e^{-\rho t} \left[ \left( \sup_{c \in L^2(M; \mathbb{R}^+)} \left\{ -\langle c, \mathcal{D}v(\tilde{k}(t)) \rangle + \langle \mathbb{1}, U(c) \rangle \right\} \right) - \left( -\langle \tilde{c}(t), \mathcal{D}v(\tilde{k}(t)) \rangle + \langle \mathbb{1}, U(\tilde{c}(t)) \rangle \right) \right] dt \geq 0 \end{aligned}$$

where we used in last step the fact that  $v$  is a solution of (21). The last expression gives  $v(k_0) - J(k_0, \tilde{c}) \geq 0$  and from the same expression we can also determine that  $v(k_0) - J(k_0, c^*) = 0$  (indeed  $c^*(\cdot)$  is defined using the feedback defined in (23)). So, for all admissible  $\tilde{c}$ ,  $v(k_0) - J(k_0, \tilde{c}) \geq 0 = v(k_0) - J(k_0, c^*)$  and then  $J(k_0, \tilde{c}) \leq J(k_0, c^*)$  and then  $c^*$  is optimal. In particular, since  $v(k_0) = J(k_0, c^*) = 0$  and  $c^*$  is an optimal control,  $v(k_0)$  is the value function at  $k_0$ . This concludes the proof.  $\square$

*Proof of Proposition 3.2.* It is part of the statement of Proposition A.8, once we read  $\langle k_0, \mathbb{1} \rangle^{1-\sigma}$  as  $(\int_M k_0(x) dx)^{1-\sigma}$ .  $\square$

*Proof of Theorem 3.3.* It is a corollary of Proposition A.8. Indeed we have proven that  $\Phi$  is an optimal feedback so the capital along the optimal trajectory is the solution of the following equation

$$\begin{cases} \dot{k}(t) = Gk(t) + Ak(t) - \Phi(k(t)) \\ k(0) = k_0 \end{cases}$$

that, using (23), is given by

$$\begin{cases} \dot{k}(t) = Gk(t) + Ak(t) - \left( \frac{\rho - A(1-\sigma)}{\sigma \operatorname{vol}(M)} \right) \langle k(t), \mathbb{1} \rangle \mathbb{1} \\ k(0) = k_0, \end{cases}$$

that is exactly (10).  $\square$

*Proof of Proposition 3.4.* From (23), we have  $c(t) = \left( \frac{\rho - A(1-\sigma)}{\sigma \operatorname{vol}(M)} \right) \langle k(t), \mathbb{1} \rangle \mathbb{1}$  that is

$$(26) \quad c^*(t, x) = \frac{\rho - A(1-\sigma)}{\operatorname{vol}(M)\sigma} K^*(t),$$

so the aggregate consumption on optimal trajectory is

$$(27) \quad C^*(t) = \langle c^*(t), \mathbb{1} \rangle = \frac{\rho - A(1-\sigma)}{\sigma} K^*(t).$$

Using such expression in (6) we have

$$\dot{K}^*(t) = AK^*(t) - \frac{\rho - A(1-\sigma)}{\sigma} K^*(t)$$

and then  $K^*(t) = K(0)e^{A - \frac{\rho - A(1-\sigma)}{\sigma}} = K(0)e^{\beta t}$  so using again respectively (27) and (26) we have  $C^*(t) = \frac{\rho - A(1-\sigma)}{\sigma} K(0)e^{\beta t}$  and  $c^*(t, x) = \frac{\rho - A(1-\sigma)}{\text{vol}(M)\sigma} K(0)e^{\beta t}$ . This concludes the proof.  $\square$

### A.3. Proof of Theorem 3.5.

**Notation A.9** The word ‘‘convergence’’ can be confusing. In this paper we always use the word as in the economic growth literature: we have convergence if the (detrended) spatial distribution of the capital tends, in the long run, to equalize across spatial locations. In mathematical terms however, we could say that  $k_D(t)(x)$  ‘‘converges’’ to a certain  $l(x)$  (when *e.g.*  $t \rightarrow \infty$ ) even if  $l(x)$  is non-constant. To avoid this possible confusion we use the expression ‘‘tends to’’ instead of ‘‘converges to’’ for this second meaning.

Theorem 3.5 is a direct consequence of the following, more detailed, results.

**Theorem A.10** *Assume that (18) is satisfied and consider an initial datum  $k_0 \in \Omega$  (defined in Proposition A.3). Then:*

1. *If  $\rho < A(1 - \sigma) + \sigma\lambda_1$  then*

$$\lim_{t \rightarrow \infty} e^{-\beta t} k(t) = \frac{\int_M k_0(x) dx}{\text{vol}(M)}, \quad \text{in } L^2(M).$$

2. *If  $\rho = A(1 - \sigma) + \sigma\lambda_1$  then (a part for a set of initial data  $k_0$  spanning a subspace of  $L^2(M)$  of co-dimension 1 i.e. a part for a ‘‘small’’ set of initial data)*

$$\lim_{t \rightarrow \infty} e^{-\beta t} k(t) = \frac{\int_M k_0(x) dx}{\text{vol}(M)} + \psi_1, \quad \text{in } L^2(M),$$

*where  $\psi_1$  is an eigenfunction related to the first non-zero eigenvalue  $\lambda_1$  of the Laplacian.*

3. *If  $\rho > A(1 - \sigma) + \sigma\lambda_1$  then (a part for a set of initial data  $k_0$  spanning a subspace of  $L^2(M)$  of co-dimension 1), we have*

$$\lim_{t \rightarrow \infty} e^{-(A-\lambda_1)t} k(t) = \psi_1, \quad \text{in } L^2(M),$$

*where  $\psi_1$  is an eigenfunction related to the first non-zero eigenvalue  $\lambda_1$  of the Laplacian.*

**Remark A.11** In Case 1 particularly, the spatial capital distribution, detrended by the factor  $e^{-\beta t}$ , tends to a spatially constant distribution. Though in Cases 2 and 3 the spatial capital distribution, detrended respectively by  $e^{-\beta t}$  and  $e^{-(A-\lambda_1)t}$ , tends to a spatially non-constant distribution. Indeed, as already recalled, any eigenfunction  $\psi_n$  related to some non-zero eigenvalue  $\lambda_n$  are non-constant. In particular, since  $\lambda_1 > 0$ , the functions  $\psi_1$  appearing in Cases 2 and 3 of Theorem A.10 are non-constant as function of the space variable. In other words, in Cases 2 and 3 we have capital agglomeration.

*Proof of Theorem A.10.* Using the feedback relation (23) into the mild form (16) we have, along the optimal trajectory,

$$k(t) = e^{\tilde{G}t} k_0 - \frac{\rho - A(1 - \sigma)}{\sigma} \int_0^t e^{(t-s)\tilde{G}} \langle k(s), \mathbb{1} \rangle \mathbb{1} ds.$$

For a given eigenfunction  $\phi_n^j$  associated to some eigenvalue  $\lambda_n$  of the Laplacian, we get

$$\begin{aligned} (28) \quad \langle k(t), \phi_n^j \rangle &= \langle e^{\tilde{G}t} k_0, \phi_n^j \rangle - \frac{\rho - A(1 - \sigma)}{\sigma \text{vol}(M)} \int_0^t \langle k(s), \mathbb{1} \rangle e^{(t-s)\tilde{G}} \langle \mathbb{1}, \phi_n^j \rangle ds \\ &= \langle k_0, e^{\tilde{G}t} \phi_n^j \rangle - \frac{\rho - A(1 - \sigma)}{\sigma \text{vol}(M)} \int_0^t \langle k(s), \mathbb{1} \rangle \langle \mathbb{1}, e^{(t-s)\tilde{G}} \phi_n^j \rangle ds. \end{aligned}$$

Using that  $e^{\tilde{G}t} \phi_n^j = e^{(-\lambda_n + A)t} \phi_n^j$  and that, for  $n \neq 0$ ,  $\langle \phi_n^j, \mathbb{1} \rangle = 0$ , we can see that

$$(29) \quad \langle k(t), \phi_n^j \rangle = \langle k_0, e^{(A-\lambda_n)t} \phi_n^j \rangle = e^{(A-\lambda_n)t} \langle k_0, \phi_n^j \rangle, \quad n \neq 0,$$

while, if  $n = 0$ , we get from Proposition 3.1 that

$$(30) \quad \langle k(t), \phi_0 \rangle = \left\langle k(t), \frac{1}{\sqrt{\text{vol}(M)}} \mathbb{1} \right\rangle = \frac{K(t)}{\sqrt{\text{vol}(M)}} = e^{\beta t} \frac{K(0)}{\sqrt{\text{vol}(M)}}.$$

*Case 1:*

We prove here the first statement, so we assume that  $\rho < A(1 - \sigma) + \sigma\lambda_1$ . Thank to this hypothesis and (18) we can fix a certain  $\varepsilon \in \left(0, \frac{A(1-\sigma) + \lambda_1 \sigma - \rho}{\sigma}\right)$ . We want to prove that  $e^{-\beta t} k(t)$  tends in the  $L^2$ -norm to  $\frac{K(0)\mathbb{1}}{\text{vol}(M)}$ .

Using (29) and (30) we have

$$\begin{aligned} \left| e^{-\beta t} k(t) - \frac{K(0)\mathbb{1}}{\text{vol}(M)} \right|_{L^2(M)}^2 &= \sum_{n \geq 0} \sum_{j=1}^{\theta_n} \left\langle \left( e^{-\beta t} k(t) - \frac{K(0)\mathbb{1}}{\text{vol}(M)} \right), \phi_n^j \right\rangle^2 = \sum_{n \geq 1} \sum_{j=1}^{\theta_n} e^{2(A-\lambda_n-\beta)t} \left\langle k_0, \phi_n^j \right\rangle^2 \\ &= e^{-2\epsilon t} \sum_{n \geq 1} \sum_{j=1}^{\theta_n} e^{2(A-\lambda_n-\beta+\epsilon)t} \left\langle k_0, \phi_n^j \right\rangle^2 \leq e^{-2\epsilon t} \sum_{n \geq 1} \sum_{j=1}^{\theta_n} \left\langle k_0, \phi_n^j \right\rangle^2 \leq e^{-2\epsilon t} |k_0|_{L^2(M)}^2 \xrightarrow{t \rightarrow \infty} 0, \end{aligned}$$

where we used that, for all  $n \neq 0$ ,  $(A - \lambda_n - \beta + \epsilon) = \left( \frac{\sigma(A-\lambda_n)-(A-\rho)+\epsilon\sigma}{\sigma} \right) < \left( \frac{\sigma(A-\lambda_1)-(A-\rho)+\epsilon\sigma}{\sigma} \right) < 0$ .

*Case 2.:*

We analyze now the case in which  $\rho = A(1 - \sigma) + \sigma\lambda_1$  so that

$$(31) \quad A - \lambda_1 - \beta = 0.$$

We introduce  $\psi_1 := \sum_{i=1}^{\theta_1} \langle k_0, \phi_1^i \rangle \phi_1^i$ .  $\psi_1$  is non-zero for all  $k_0 \in \Omega$  except those contained in a subspace of co-dimension  $\theta_1 \geq 1$  of  $L^2(M)$ . To prove that  $e^{-\beta t} k(t)$  tends to  $\frac{K(0)\mathbb{1}}{\text{vol}(M)} + \psi_1$  in the  $L^2$ -norm we first observe that

$$(32) \quad \left\langle \psi_1, \phi_n^j \right\rangle = 0, \quad \text{for any } n \neq 1$$

and

$$(33) \quad \left\langle \psi, \phi_1^j \right\rangle = \left\langle k_0, \phi_1^j \right\rangle, \quad \text{for any } j \in \{1, \dots, \theta_1\}.$$

Using these facts together with (29), (30) and (31) we have

$$\begin{aligned} \left| e^{-\beta t} k(t) - \psi_1 - \frac{K(0)\mathbb{1}}{\text{vol}(M)} \right|_{L^2(M)}^2 &= \sum_{n \geq 0} \sum_{j=1}^{\theta_n} \left\langle \left( e^{-\beta t} k(t) - \psi_1 - \frac{K(0)\mathbb{1}}{\text{vol}(M)} \right), \phi_n^j \right\rangle^2 \\ &= \left( \frac{K(0)}{\sqrt{\text{vol}(M)}} - \langle \psi_1, \phi_0 \rangle - \left\langle \frac{K(0)\mathbb{1}}{\text{vol}(M)}, \phi_0 \right\rangle \right)^2 + \left( \sum_{j=1}^{\theta_1} e^{(A-\lambda_1-\beta)t} \left\langle k_0, \phi_1^j \right\rangle - \langle \psi_1, \phi_1^j \rangle - \left\langle \frac{K(0)\mathbb{1}}{\text{vol}(M)}, \phi_1^j \right\rangle \right)^2 \\ &+ \sum_{n \geq 2} \sum_{j=1}^{\theta_n} e^{2(A-\lambda_n-\beta)t} \left\langle k_0, \phi_n^j \right\rangle^2 = 0 + \left( \sum_{j=1}^{\theta_1} \left\langle k_0, \phi_1^j \right\rangle - \langle \psi_1, \phi_1^j \rangle - 0 \right)^2 + \sum_{n \geq 2} \sum_{j=1}^{\theta_n} e^{2(A-\lambda_n-\beta)t} \left\langle k_0, \phi_n^j \right\rangle^2 \\ &= 0 + 0 + \sum_{n \geq 2} \sum_{j=1}^{\theta_n} e^{2(A-\lambda_n-\beta)t} \left\langle k_0, \phi_n^j \right\rangle^2. \end{aligned}$$

We fix now  $\epsilon \in (0, \lambda_2 - \lambda_1)$ . The previous expression equals

$$e^{-2\epsilon t} \sum_{n \geq 2} \sum_{j=1}^{\theta_n} e^{2(A-\lambda_n-\beta+\epsilon)t} \left\langle k_0, \phi_n^j \right\rangle^2 \leq e^{-2\epsilon t} \sum_{n \geq 2} \sum_{j=1}^{\theta_n} \left\langle k_0, \phi_n^j \right\rangle^2 \leq e^{-2\epsilon t} |k_0|_{L^2(M)}^2 \xrightarrow{t \rightarrow \infty} 0,$$

where in the first inequality we used that for all  $n \geq 2$ , thanks to (11) and (31),  $(A - \lambda_n - \beta + \epsilon) \leq (A - \lambda_2 - \beta + \epsilon) < (A - \lambda_2 - \beta + (\lambda_2 - \lambda_1)) = 0$ . The last limit holds because  $\epsilon > 0$ .

*Case 3.:*

The condition  $\rho > A(1 - \sigma) + \sigma\lambda_1$ , is equivalent to  $(A - \lambda_1 + \beta) > 0$ .  $\psi_1$  appearing in the text of the proposition is, as in case 2, given by  $\sum_{i=1}^{\theta_1} \langle k_0, \phi_1^i \rangle \phi_1^i$ . Thanks to (30), (32) and then (29) we have

$$\begin{aligned} \left| e^{-(A-\lambda_1)t} k(t) - \psi_1 \right|_{L^2(M)}^2 &= \sum_{n \geq 0} \sum_{j=1}^{\theta_n} \left\langle \left( e^{-(A-\lambda_1)t} k(t) - \psi_1 \right), \phi_n^j \right\rangle^2 \\ &= \left( e^{-(A-\lambda_1+\beta)t} \frac{K(0)}{\sqrt{\text{vol}(M)}} + 0 \right)^2 + \left( \sum_{j=1}^{\theta_1} \left\langle k_0, \phi_1^j \right\rangle - \langle \psi_1, \phi_1^j \rangle \right)^2 + \sum_{n \geq 2} \sum_{j=1}^{\theta_n} e^{2(A-\lambda_n-(A-\lambda_1))t} \left\langle k_0, \phi_n^j \right\rangle^2 \end{aligned}$$



that, thanks to (33) equals

$$\begin{aligned}
(34) \quad & e^{-2(A-\lambda_1+\beta)t} \left( \frac{K(0)}{\sqrt{\text{vol}(M)}} \right)^2 + 0 + \sum_{n \geq 2} \sum_{j=1}^{\theta_n} e^{2(\lambda_1-\lambda_n)t} \langle k_0, \phi_n^j \rangle^2 \\
& \leq e^{-2(A-\lambda_1+\beta)t} \left( \frac{K(0)}{\sqrt{\text{vol}(M)}} \right)^2 + e^{2(\lambda_1-\lambda_2)t} \sum_{n \geq 2} \sum_{j=1}^{\theta_n} \langle k_0, \phi_n^j \rangle^2 \\
& \leq e^{-2(A-\lambda_1+\beta)t} \left( \frac{K(0)}{\sqrt{\text{vol}(M)}} \right)^2 + e^{2(\lambda_1-\lambda_2)t} |k_0|_{L^2(M)}^2 \xrightarrow{t \rightarrow \infty} 0
\end{aligned}$$

where we used (11) in first inequality and we concluded using that  $(A - \lambda_1 + \beta) > 0$  and that  $\lambda_1 < \lambda_2$ . This concludes the proof.  $\square$

**Remark A.12** In the proof of Theorem A.10 we have shown that the limits hold in the  $L^2(H)$  sense but with a similar argument we could show that in fact the limits are uniform.