

Long-memory process and aggregation of AR(1) stochastic processes: A new characterization

Bernard Candelpergher, Michel Miniconi, Florian Pelgrin

► **To cite this version:**

| Bernard Candelpergher, Michel Miniconi, Florian Pelgrin. Long-memory process and aggregation of AR(1) stochastic processes: A new characterization. 2015. <hal-01166527v2>

HAL Id: hal-01166527

<https://hal.univ-cotedazur.fr/hal-01166527v2>

Submitted on 7 Aug 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.

Long-memory process and aggregation of AR(1) stochastic processes: A new characterization

Bernard Candelpergher^a, Michel Miniconi^b and Florian Pelgrin^c

Abstract

Contemporaneous aggregation of individual AR(1) random processes might lead to different properties of the limit aggregated time series, in particular, long memory (Granger, 1980). We provide a new characterization of the series of autoregressive coefficients, which is defined from the Wold representation of the limit of the aggregate stochastic process, in the presence of long-memory features. Especially the infinite autoregressive stochastic process defined by the almost sure representation of the aggregate process has a unit root in the presence of the long-memory property. Finally we discuss some examples using some well-known probability density functions of the autoregressive random parameter in the aggregation literature.

JEL Classification Code: C2, C13.

Keywords: Autoregressive process, Aggregation, Heterogeneity, Complex variable analysis.

^aUniversity of Nice Sophia-Antipolis, Laboratoire Jean-Alexandre Dieudonné. E-mail: bernard.candelpergher@unice.fr.

^bUniversity of Nice Sophia-Antipolis, Laboratoire Jean-Alexandre Dieudonné. E-mail: michel.miniconi@unice.fr.

^cEDHEC Business School, E-mail: florian.pelgrin@edhec.edu.

This paper has benefited from very useful comments from and discussions with Stéphane Gregoir. The usual disclaimer applies.

1 Introduction

Aggregation is a critical and widely acknowledged issue in the empirical and theoretical literature in economics and other fields. Especially, since the contributions of Granger and Morris (1976) and Granger (1980), it is well-known that the contemporaneous aggregation of individual random AR(1) stochastic processes might lead to long memory models.¹ Notably, Granger (1980) considers the case of a Beta-distribution for the random autoregressive parameter and thus points out that the long memory property depends on the behavior of the density of the random autoregressive parameter near unity and that the common and idiosyncratic components might exhibit a different degree of long memory. In particular, Zaffaroni (2004) generalizes these results by studying the limit of the aggregate process with a quite flexible (semi-parametric) assumption regarding the behavior near unity of the probability density function of the random autoregressive coefficient and makes clear the asymptotics of both the common and idiosyncratic parts of the aggregated process. Among others results, Zaffaroni (2004) shows formally that the more concentrated is the distribution of the random autoregressive coefficient near the unit, the stronger is the long-memory property of the limit aggregated process.² Following these contributions, we study the aggregation of heterogenous individual AR(1) stochastic processes that leads to the long-memory property. In contrast to the literature, we focus on the infinite autoregressive representation of the limit of the aggregate process, and especially the sum of the autoregressive coefficients, rather than the usual infinite moving representation.

Indeed the use of the (equivalent) infinite autoregressive representation of the limit of the aggregated process might be meaningful in various contexts of the aggregation (panel) literature. As pointed out by Lewbel (1994), the autoregressive representation is insightful among others when estimating the (macro) aggregate dynamics with unobservable individual series in the presence of unobserved parameter heterogeneity, when identifying and estimating certain distributional features of the micro parameters from aggregate re-

¹In the sequel, a stochastic process is said to have a long memory property if its autocovariance function is not summable (Beran, 1994; Beran et al., 2013).

²For a recent survey, see Leipus et al. (2013).

lations (disaggregation problem), or when deriving the aggregate dynamics with unknown common error terms.³ Obviously, if the limit of the aggregate process is represented almost surely by a short memory process, the infinite autoregressive representation and especially the determination of its persistence (e.g., through the sum of the autoregressive coefficients) easily obtains using standard results of time series analysis (Brockwell and Davis, 2002). In contrast, if the limit of the aggregate process is represented almost surely by a long memory process, the convergence and thus the determination of the series of autoregressive coefficients is challenging since the series of moving average coefficients is no longer absolutely summable. Such a characterization might be necessary for instance in the case of the estimation of the aggregate dynamics with unobservable individual series in which finite parameter approximation for the infinite lag distribution is required—the autoregressive representation of the limit aggregated process displays an infinite parametrization whereas the econometrician has only finitely many observations and thus finite-parameter approximations might lead to near observational aggregate processes with different statistical properties as long as the sum of the autoregressive coefficients is not correctly identified. In this respect, our paper tackles this issue and proposes a new characterization of the autoregressive representation of limit long-memory aggregate processes that results from individual random AR(1) stochastic processes.

Notably aggregation of individual random AR(1) processes is analyzed in this paper under the assumptions that some common factors exist, the (positive) random autoregressive coefficient takes values in $[0, 1)$ with probability distribution μ , and all of the noncentral moments exist. In particular, the series of all of the noncentral moments might be either absolutely summable or divergent and the limit of the aggregated stochastic process satisfies almost surely both an infinite moving average and autoregressive representations. Within this framework, we show that the sum of the autoregressive coefficients of the stochastic process defined from the limit in \mathcal{L}^2 of the aggregate process, say $\sum_{k \geq 1} a_k$, might equal one (respectively, less than one) when the limit aggregate process has the long memory property (respectively, short memory property). Say differently, the divergence of the series of noncentral moments is fully equivalent to the presence of a unit root in

³See also Pesaran (2003), Pesaran and Chudik (2014), Jondeau and Pelgrin (2014a,b).

the representation of the stochastic process defined by the limit of the aggregate process. In so doing, we consider the mapping between the moving average coefficients and the autoregressive coefficients of the limit aggregate process and make use of complex analysis when the series of the moving average coefficients is divergent. Complex analysis is called for at least two arguments. On the one hand, the existence of the limit $\lim_{r \rightarrow 1^-} a(r)$ of the (ordinary) generating function of the sequence of the autoregressive coefficients does not insure its equality to $\sum_{k \geq 1} a_k$ or even the convergence of this sum. On the other hand, we cannot apply a standard Hardy-Littlewood Tauberian theorem since, to the best of our knowledge, there is no general proof of the positiveness of the autoregressive coefficients for all probability distributions μ defined on $[0, 1)$. Interestingly such a complex analysis might be used to study the behavior on the unit circle of an infinite parameterized stochastic process, which is defined (almost surely) in \mathcal{L}^2 but does not belong to \mathcal{L}^1 .

The rest of the paper is organized as follows. In Section 2, we discuss the main assumptions regarding the individual random AR(1) stochastic processes and then we derive the limit aggregate process. In Section 3, we provide the two main results of our paper, namely the determination of the sum of the autoregressive parameters and the characterization of the stochastic process defined by the limit in \mathcal{L}^2 of the aggregated process. In Section 4, we assume that the distribution of the random autoregressive coefficient is subsequently a Beta distribution (of the first kind), a uniform distribution and a random polynomial density function. In so doing, we apply the results of Section 3 and generalizes some results in the literature. Proofs are gathered in the Appendix.

2 Aggregating individual AR(1) processes

In this section, we first discuss the assumptions regarding the individual AR(1) processes. Then we derive the limit aggregate process.

2.1 Individual dynamics

Consider the individual random AR(1) model for $i = 1, \dots, N$ and $t \in \mathbb{Z}$:⁴

$$x_{i,t} = \varphi_i x_{i,t-1} + v_{i,t}, \quad (1)$$

where φ_i denotes an individual-specific (random) parameter drawn from a fixed random variable φ and $v_{i,t}$ is an error term that can be decomposed into a common component, ϵ_t , and an idiosyncratic (individual-specific) component, $\eta_{i,t}$.⁵

$$v_{i,t} = \epsilon_t + \eta_{i,t}. \quad (2)$$

The macro variable results from the aggregation of micro-units, with the use of time-invariant nonrandom weights $W_N = (w_1, \dots, w_N)'$, with $\sum_{i=1}^N w_i = 1$, so that the aggregate variable is defined as the weighted average of the micro-units $X_{N,t} = \sum_{i=1}^N w_i x_{i,t}$. The following assumptions hold:

Assumption 1: φ a fixed random variable with probability distribution μ the support of which is in $[0, 1]$.

Assumption 2: The moments $u_k = \mathbb{E}(\varphi^k)$ exist for all integer $k \geq 1$.

Assumption 3: ϵ_t and $\eta_{i,t}$ are white noise processes with means of zero and variance of σ_ϵ^2 and σ_η^2 , respectively; ϵ_t and $\eta_{i,t}$ are mutually orthogonal at any lag and lead.

Assumption 4: Realizations of φ are independent of those of ϵ_t .

Assumption 5: As $N \rightarrow \infty$, $\|W_N\| = O(N^{-1/2})$ and $w_i / \|W_N\| = O(N^{-1/2})$ for all $i \in \mathbb{N}$.

Before discussing our main assumptions, we introduce some notations. Let \mathcal{H}_x denote the Hilbert space generated by all of the random variables that compose the panel

⁴Alternatively one might assume that $t \in \mathbb{N}$ and consider an asymptotically stationary limit aggregate process (Gonçalves and Gourieroux, 1994).

⁵Such dynamics have been used in economics (among others) to represent consumption expenditures across households (Lewbel, 1994), consumer price inflation across subindices (Altissimo et al., 2009), real exchange rates across sectors (Imbs et al., 2005), or real marginal cost across industries (Imbs et al., 2011).

$(x_{i,t})_{i=1,\dots,N,t\in\mathbb{Z}}$, \mathcal{H}_{x_i} the Hilbert space generated by the stochastic process $(x_{i,t})_{t\in\mathbb{Z}}$, $\mathcal{H}_{x,t}$ and $\mathcal{H}_{x_i,t}$ the corresponding subspaces of \mathcal{H}_x and \mathcal{H}_{x_i} up to time t . Assumption 1 outlines that φ is a random variable with distribution on the support $[0, 1)$ such that $\varphi \in \bigcap_t \bigoplus_i \mathcal{H}_{x_i,t}$.⁶ This assumption is consistent with many parametric specifications of the cross-sectional distribution of φ . We only rule out situations in which some individual processes are not almost surely (asymptotically) stationary, i.e. $\mathbb{P}(|\varphi| \geq 1) > 0$ (see further). In particular, Assumption 1 includes the Beta distribution, $B(p, q)$. In this case, the representation of the (limit) aggregate process does depend on the mass distribution of the Beta distribution around unity: the smaller is q , the larger is the mass of the distribution around unity (Zaffaroni, 2004). In contrast, imposing the condition $0 \leq \varphi \leq c < 1$ for some constant c will guarantee that there are no individual unit root parameters that would dominate at the aggregate level (Zaffaroni, 2004) and that the limit aggregate process (as $N \rightarrow \infty$) displays short memory with an exponentially decaying autocorrelation function. Assumption 2 insures that noncentral moments $u_k = \mathbb{E}(\varphi^k)$ of any (nondegenerate) random variable φ , defined on $[0, 1)$, satisfy: $1 > u_1 \geq \dots \geq u_k \geq 0, \forall k \geq 1$, and $u_k \rightarrow 0$ as $k \rightarrow \infty$.

Assumption 3 defines the statistical properties of the two components of the error term, the common shock (factor) and the idiosyncratic shock. Several points are worth commenting. First, without loss of generality, one might assume that the stochastic process $(v_{i,t})_{i=1,\dots,N,t\in\mathbb{Z}}$ is weakly linearly exchangeable (Aldous, 1980), i.e. $(v_{i,t})_{i=1,\dots,N,t\in\mathbb{Z}}$ is a purely non deterministic (regular) covariance stationary process and the covariance structure is invariant by translation with respect to the time index and invariant by permutation with respect to the unit index. Second, the stochastic process $(\epsilon_t)_{t\in\mathbb{Z}}$ is assumed to be known and given.⁷ Given that $\mathcal{H}_x = \bigoplus_i \mathcal{H}_{x_i}$ and thus $\mathcal{H}_{x,t} = \bigoplus_i \mathcal{H}_{x_i,t}$, one has $\epsilon_t \in \mathcal{H}_{x,t} \cap (\mathcal{H}_{x,t-1})^\perp$. Third, taking that $\mathcal{H}_{x_i} = \mathcal{H}_\epsilon \oplus \mathcal{H}_{x_i-\epsilon}$, the idiosyncratic stochastic process is such that $\eta_{i,t} \in \mathcal{H}_{x_i-\epsilon,t} \cap (\mathcal{H}_{x_i-\epsilon,t-1})^\perp$. Fourth, the assumption that the vari-

⁶Note that the set $(-1, 0)$ has been excluded from the support of φ for simplicity's sake.

⁷Without altering our main results but at the expense of further assumptions, (i) the common error term at time t might be multiplied by the realization of a (scaling) random variable κ to introduce some form of heteroscedasticity, (ii) the stochastic process $(\epsilon_t)_{t\in\mathbb{Z}}$ might be assumed to be unknown (using a two-step procedure), and (iii) multiple independent common error terms might be introduced in the specification (Zaffaroni, 2004).

ance of the idiosyncratic shocks is the same across individuals might be too stringent in practical examples (e.g., in economics or in finance) and can be generalized by assuming heteroscedastic idiosyncratic error terms. Eq. (2) together with Assumption 3 provide a parsimonious form of (statistical) cross-sectional dependence, which is common in the aggregation literature (Forni and Lippi, 2001; Zaffaroni, 2004). Obviously one may introduce dependence on the idiosyncratic part through cross-correlation and/or spatial correlation (Chudik, Pesaran and Tosetti, 2011; Pesaran and Tosetti, 2009). Such correlation has no effect on our results as long as it is sufficiently weak (e.g., local correlation) so that the error structure belongs to the class of approximate factor structure models (Chamberlain and Rothschild, 1983).⁸ Assumption 4 avoids any correlation between the error terms and φ . Assumption 5 is a granularity condition, which insures that the weights used to define the aggregate process are not dominated by a few of the cross-sectional units (Gabaix, 2011; Pesaran and Chudik, 2014).⁹

2.2 Aggregate Dynamics

The empirical cross-sectional moments of φ are $\tilde{\mathbb{E}}_N(\varphi^k) = \sum_{i=1}^N w_i \varphi_i^k, \forall k \geq 1$. For sake of simplicity and without loss of generalization with respect to Assumption 4, we assume that $w_i = 1/N$ for all i . Consequently, as $N \rightarrow \infty$, $\tilde{\mathbb{E}}_N(\varphi^k) \xrightarrow{a.s.} u_k$.

Using Eqs. (1)–(2), the exact aggregate dynamics can be written as:¹⁰

$$\prod_{j=1}^N (1 - \varphi_j L) X_{N,t} = \frac{1}{N} \sum_{i=1}^N \prod_{j \neq i} (1 - \varphi_j L) v_{i,t} \quad (3)$$

or equivalently

$$X_{N,t} = \frac{1}{N} \sum_{i=1}^N (1 - \varphi_i L)^{-1} \epsilon_t + \frac{1}{N} \sum_{i=1}^N (1 - \varphi_i L)^{-1} \eta_{i,t} \quad (4)$$

⁸The condition is that the maximum eigenvalue of the covariance matrix of $\eta_t = (\eta_{1,t}, \dots, \eta_{N,t})'$ remains bounded as the cross-section dimension increases.

⁹Our results extend to the case of (time-varying) stochastic weights. Such an extension requires at least that the weights be distributed independently from the stochastic process defining the random variable.

¹⁰Put differently, it is an ARMA(N,N-1) in the absence of common roots in the individual processes (Granger and Morris, 1976).

where L is the lag operator ($z_{t-1} = Lz_t$). Taking Eq. (4), we can characterize the asymptotic behavior of both the idiosyncratic component and the common component. This is done in the following proposition. Results are known but are reported here for sake of completeness.

Proposition 1 *Suppose that Assumptions 1–5 hold. Given the disaggregate model defined in Eqs. (1)–(2), the limit in \mathcal{L}^2 of the aggregated process as $N \rightarrow \infty$ satisfies (almost surely) the two equivalent representations:*

$$X_t = \sum_{k=0}^{\infty} u_k \epsilon_{t-k} \quad (MA \text{ form}), \quad (5)$$

$$X_t = \sum_{k=1}^{\infty} a_k X_{t-k} + \epsilon_t \quad (AR \text{ form}), \quad (6)$$

where $X_{N,t} \xrightarrow{\mathcal{L}^2} X_t$ and $\tilde{\mathbb{E}}_N(\varphi^k) \xrightarrow{a.s.} u_k = \mathbb{E}(\varphi^k)$ as $N \rightarrow \infty$. The sequence $\{a_k, k \geq 1\}$ where $a_k = \mathbb{E}[A_k]$ satisfies the recurrence relation :

$$A_1 = \varphi, \quad A_{k+1} = (A_k - a_k) \varphi \quad (7)$$

Proof: Gonçalves and Gouriéroux (1988), Lewbel (1994).

Several points are worth commenting. First, as shown by Robinson (1978), with the exception of a degenerate distribution for φ (e.g., Dirac distribution), the dynamics of the limit aggregate process is richer than the individual dynamics because of the non-ergodicity of the individual random AR(1) process. Second, using the infinite moving average representation and the positiveness of the moments, the (limit) aggregate process displays short memory if $\sum_{k=0}^{\infty} u_k < +\infty$ whereas it has a long-memory property if $\sum_{k=0}^{\infty} u_k = +\infty$. It is worth noting that (i) the (limit) aggregate process is in \mathcal{L}^2 but not necessarily in \mathcal{L}^1 , and (ii) the sum of the autoregressive coefficients is (absolutely) convergent if the aggregate process has short memory (since the spectral representation is unique). Third, the central moments of the cross-sectional distribution of φ can be easily obtained from the infinite autoregressive representation of the aggregate process. This is useful when considering the standard disaggregation problem in statis-

tics. For instance, the first four cross-sectional moments are $\mathbb{E}[\varphi] = a_1$, $\mathbb{V}[\varphi] = a_2$, $S[\varphi] = (a_3 - a_1 a_2) / (a_2)^{3/2}$, and $K[\varphi] = (a_4 - 2a_1 a_3 + a_1^2 a_2 + a_2^2) / (a_2)^2$. Fourth, Equation (6) shows that aggregation leads to an infinite autoregressive model for X_t (see Robinson, 1978, Lewbel, 1994). Notably, using Eq. (7), the autoregressive parameters a_k are nonlinear transformations of the noncentral moments of φ and satisfy the following non-homogenous difference equations (for $k \geq 2$):

$$\begin{aligned} a_{k+1} &\equiv \mathbb{E}[A_{k+1}] = u_{k+1} - \sum_{r=1}^k a_r u_{k-r+1} \\ a_1 &= \mathbb{E}[\varphi]. \end{aligned}$$

Fifth, the persistence of the stochastic process defined by the limit of the aggregated process can be defined as the sum of the autoregressive coefficients, which is denoted $a(1) = \sum_{k=1}^{\infty} a_k$. More specifically, if the limit of the aggregate process belongs to \mathcal{L}^1 and thus $(X_t)_{t \in \mathbb{Z}}$ is a short-memory stochastic process, then it is straightforward to show that

$$a(1) = \left(\mathbb{E} \left[\frac{1}{1-\varphi} \right] \right)^{-1} \mathbb{E} \left[\frac{\varphi}{1-\varphi} \right].$$

or equivalently

$$(1 - a(1))^{-1} = \mathbb{E} \left[\frac{1}{1-\varphi} \right].$$

In particular, the limit of the aggregate stochastic process has short memory if and only if $\mathbb{E} \left[\frac{1}{1-\varphi} \right] < \infty$ (Robinson, 1978; Gonçalves and Gouieroux, 1994). In the spirit of Zaffaroni (2004), a sufficient condition on the probability density function h of φ for the short memory property is that there exists $\alpha \in (0, 1)$ and a constant C such that $\lim_{x \rightarrow 1^-} \frac{h(x)}{(1-x)^{1-\alpha}} = C$. Sixth, we can define the generating function of the a_k terms as follows. Taking the recurrence relation, one can write formally

$$\sum_{k \geq 1} A_k z^k = \frac{z\varphi}{1 - z\varphi} \left(1 - \sum_{k \geq 1} a_k z^k \right)$$

and thus

$$\sum_{k \geq 1} a_k z^k = \sum_{k \geq 1} u_k z^k - \sum_{k \geq 1} a_k z^k \sum_{k \geq 1} u_k z^k.$$

Let m denote

$$m(z) = \sum_{k \geq 1} u_k z^k,$$

we obtain the formal generating function of the a_k terms

$$a(z) \equiv \sum_{k \geq 1} a_k z^k = \frac{m(z)}{1 + m(z)}, \quad (8)$$

and thus the mapping between the generating function of the infinite autoregressive lag polynomial and the one of the infinite moving average representation. As explained in Section 3, Eq. (8) is fundamental since our proof is intimately related to the function $\frac{m}{1+m}$.

3 Aggregate long-memory process and the sum of the autoregressive parameters

In this section, we show that the sum of the autoregressive coefficients equals one in the presence of a (limit) aggregate process in \mathcal{L}^2 but not in \mathcal{L}^1 . In so doing, we emphasize that complex analysis is required to obtain the convergence of the series $\sum_{k \geq 1} a_k$. Then we provide a characterization of the (limit) aggregate long-memory process in the presence of individual random AR(1) processes.

3.1 The function m

Before showing the convergence of the series of autoregressive parameters, we need an intermediate result regarding the m function, and especially to make clear the relationship between the sum of the moving average coefficients $\sum_{k \geq 1} u_k$ and $m(r)$ when $r \rightarrow 1^-$. Proposition 2 clarifies this link and turns to be extremely useful when characterizing

the infinite sum of the autoregressive coefficients of a (limit) aggregate long-memory process.¹¹

Proposition 2 $\sum_{k=1}^{\infty} u_k = +\infty$ if and only if $\lim_{r \rightarrow 1^-} m(r) = +\infty$

Proof : See Appendix A.

3.2 The series $a(z) = \frac{m(z)}{1+m(z)}$ and the convergence of $\sum_{k \geq 1} a_k$

One key issue to study the convergence of the autoregressive coefficients is that the existence of the limit, $\lim_{r \rightarrow 1^-} a(r)$, does not insure its equality to $\sum_{k \geq 1} a_k$ or even the convergence of this sum.¹² Supplementary *Tauberian* conditions are needed for this kind of results (Hardy, 1949; Titchmarsh, 1939; Korevaar, 2004). In particular, the convergence of this series has to be studied by making use of complex analysis, especially in the case of a (limit) aggregate long-memory process.

All of the proofs related to Theorem 1 are gathered in Appendix B. To summarize, the proof of the convergence of the autoregressive coefficients proceeds as follows. We first define m as an analytic function of the complex variable z within the open disc around 0 of radius one, $D(0, 1)$, and rewrite $m(z)$ as an integral on $[0, 1)$ with parameter $z \in D(0, 1)$. This allows showing that it can be continuously extended to $\overline{D(0, 1)} \setminus \{1\}$ (see Lemma 1 in Appendix B). Then a second lemma (see Lemma 2 in Appendix B) proves that the function $a(z) = \frac{m(z)}{1+m(z)}$ is well defined in the disc $D(0, 1)$ (i.e., the denominator does not vanish) and then it can be extended to the whole closed disc $\overline{D(0, 1)}$. For this purpose, two cases must be studied, according to the nature of the series $\sum u_k$: if $\sum u_k < \infty$ the function m is continuous in the closed disc, so is a ; and if $\sum u_k = \infty$ then $|m(z)| \rightarrow \infty$ when $z \rightarrow 1$ in $D(0, 1)$, therefore $a(z) \rightarrow 1$ when $z \rightarrow 1$ in $D(0, 1)$. Finally, a third lemma shows that the analytic function a is univalent and provides the

¹¹Note obviously that $\sum_{k=0}^{\infty} u_k = +\infty$ if $\lim_{r \rightarrow 1^-} m(r) = +\infty$.

¹²For instance, consider the power series defined for $|r| < 1$, $h(r) = \sum_{k=0}^{\infty} (-1)^k r^k$. Then $h(r) = \frac{1}{1+r}$ and $\lim_{r \rightarrow 1^-} h(r) = \frac{1}{2}$. However $\sum_0^{\infty} (-1)^k$ is not convergent.

use of a Tauberian condition (see Lemma 3 in Appendix B). This then allows to prove the convergence of the series $\sum a_n$ (Theorem 1).

Theorem 1 *Let $\{a_k, k \geq 1\}$ denote the sequence defined in Proposition 1. The series $\sum_{k=1}^{\infty} a_k$ is convergent and*

$$\sum_{k=1}^{+\infty} a_k = \lim_{r \rightarrow 1^-} \frac{m(r)}{1 + m(r)}$$

Proof : See Appendix B.

Taking Proposition 2 and Theorem 1, it is then straightforward to determine the sum of the autoregressive coefficients.

Proposition 3 *Let $\{a_k, k \geq 1\}$ denote the sequence defined in Proposition 1. The sum of the autoregressive coefficients, $\sum_{k=1}^{\infty} a_k$, equals one if and only if $\lim_{r \rightarrow 1^-} m(r) = +\infty$ or equivalently if and only if $\sum_{k=1}^{\infty} u_k = +\infty$.*

Three points are worth commenting. First, if the limit aggregate process is a second-order stationary process (Proposition 1) and the series of its moving average coefficients is absolutely summable, then the sum of the autoregressive coefficients is less than one. Notably, this result obtains with classical time series results whereas there is a need of complex analysis when the series of the moving average coefficients is not absolutely summable and one studies what happens at the pole $z = 1$. A second and related point is that the behavior of the series of the autoregressive coefficients depends on whether the limit second-order stationary aggregate process belongs to \mathcal{L}^1 or not. Consequently, as explained below, this provides a new characterization of a (limit) aggregate long-memory process as a result of the aggregation of random AR(1) processes. Third, as stated in Corollary 1, the function $\Phi(z) = 1 - a(z)$ admits only one zero on $\overline{D(0, 1)}$.

Corollary 1 *Let $\Phi(z) = 1 - a(z)$. Then $z = 1$ is a zero of the function Φ , which is defined on $\overline{D(0, 1)}$, if and only if $\sum_{k=1}^{\infty} u_k = +\infty$.*

Proof : This is a straightforward implication of Theorem 1.

Corollary 1 establishes that the representation of the stochastic process defined by the limit of the aggregate process $(X_{N,t})_{t \in \mathbb{Z}}$ admits a unit root whereas the aggregate process is weakly stationary. There is a one-to-one relationship between the long memory property and the presence of a unit root of the infinite autoregressive limit aggregate process.

3.3 Time series implications

Taking Theorem 1 and Proposition 3, we are now in a position to provide a new characterization of a (limit) aggregate long-memory process that results from the aggregation of individual random AR(1) processes.

Theorem 2 *Suppose that Assumptions 1 and 2 hold true. Let $(X_t, t \in \mathbb{Z})$ denote a long-memory process with the following Wold decomposition*

$$X_t = \sum_{k=0}^{\infty} u_k \epsilon_{t-k}$$

where the u_k terms are nonnegative and $\sum_{k=0}^{\infty} u_k = \infty$. Then the a_j terms of the equivalent infinite autoregressive representation,

$$X_t = \sum_{k=1}^{\infty} a_k X_{t-k} + \epsilon_t,$$

which are defined from $a_{k+1} = u_{k+1} - \sum_{r=1}^k a_r u_{k-r+1}$ and $a_1 = u_1$, satisfy $\sum_{k=1}^{\infty} a_k = 1$.

Several points are worth discussing. First, the (limit) aggregate long-memory process does not belong to the class of ARFIMA or self-similar stochastic processes.¹³ In particular, it is not possible to express the moving average coefficients of a fractional ARIMA process (by binomial expansion) such that they match the moving average weights of the

¹³A continuous-time stochastic process (Y_t) is said to be self-similar with self-similarity parameter H , if for any sequence of time points t_1, \dots, t_k and any positive constant a , $c^{-H}(Y_{at_1}, \dots, Y_{at_k})$ has the same distribution as $(Y_{t_1}, \dots, Y_{t_k})$.

long-memory process defined in Theorem 2. In the same respect, matching the aggregate moving average coefficients of the (limit) aggregate process requires non-constant values of the self-similarity parameter H (Beran, 1994; Beran et al., 2013). Second, the fact that the sum of the autoregressive coefficients of the (limit) aggregate process is equal to one is also consistent with the standard definition of long-memory processes (Beran, 1994; Beran et al., 2013), i.e. there exists a real number $\alpha \in (0, 1)$ and a constant c_ρ such that $\lim_{k \rightarrow \infty} \frac{\rho(k)}{c_\rho k^\alpha} = 1$ where $\rho(k) = \gamma_X(h)/\gamma_X(0)$ (with $\gamma_X(h) = \mathbb{E}[\varphi^h/(1 - \varphi^2)]$) is the autocorrelation of order k of the (limit) aggregate stochastic process (X_t) .

Third, as pointed out by Beran (1994), observing long-range dependence in an aggregate time series (e.g., at the macro-level) does not necessarily mean that this is due to the genuine occurrence of long memory in the individual series (micro-level). This might be induced artificially by aggregation. Say differently, identifying the source of long memory would require to look carefully at the behavior of the possibly unobservable individual series. Fourth, Theorem 2 has some implications, which are beyond the scope of this paper, regarding some aggregation problems often encountered in the theoretical and empirical aggregation research (Pesaran and Chudik, 2014), namely the derivation of the macro dynamics from heterogenous individual dynamics and the identification and estimation of certain distributional features of the micro parameters from aggregate relations (disaggregation problem).¹⁴ For instance, in the case of the disaggregation problem when individual series are not available, since the autoregressive specification in Proposition 1 displays an infinite parametrization and the econometrician has only finitely many observations, one might proceed with a finite-parameter approximation for the infinite lag distribution (Sims, 1971, 1972; Faust and Lepper, 1997; Pötscher, 2002) and might account possibly for the constraint $a(1) = 1$. Notably, finite-parameter approximations might lead to near observational aggregate processes with different statistical properties and thus to the incorrect identification of distributional features of the micro parameters.¹⁵ Other applications might concern, among others, the estimation of the aggregate

¹⁴See Lewbel (1994), and Jondeau and Pelgrin (2014b).

¹⁵The convergence of the estimates of the approximation is not sufficient to guarantee the convergence of some functions of those parameters—pointwise convergence does not imply (locally) uniform convergence.

dynamics in the presence of unobserved heterogeneity when individual series are not available, the derivation of the macrodynamics in heterogeneous agents models, or the reconciliation of micro (panel) evidence and macro facts (e.g., the purchasing power parity, the nominal price rigidity).

Finally it is worth noting that Theorem 2 applies in a broader context than the aggregation of random AR(1) stochastic processes. Indeed, any stochastic process defined (almost surely) in \mathcal{L}^2 that admits a Wold decomposition with decreasing and nonnegative moment moving average coefficients and does not belong to \mathcal{L}^1 displays a unit root. This comes from the formal identity between the generating functions of (a_k) and (u_k) , $\sum_{k \geq 1} a_k z^k = \sum_{k \geq 1} u_k z^k - \sum_{k \geq 1} a_k z^k \sum_{k \geq 1} u_k z^k$.

4 Examples

In this section, we review three examples, namely the Beta $B(p, q)$ distribution (of the first kind), the uniform distribution with $p = q = 1$, and the random polynomial density.

Beta distribution Following Morris and Granger (1976), Granger (1980) and Gonçalves and Gouriéroux (1994), we assume that φ is Beta-distributed.¹⁶

$$B(p, q; x) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} x^{p-1} (1-x)^{q-1} \mathbf{1}_{[0,1]}(x), \quad p > 0, \quad q > 0.$$

In this respect, to the best of our knowledge, Proposition 4 provides a new characterization of the series of autoregressive coefficients (i.e., the persistence of the (limit) aggregate process (X_t)).

Proposition 4 *Suppose that φ is Beta-distributed (of the first kind). Given the disaggregate model defined in Eqs. (1)–(2), the series of the autoregressive coefficients of the limit aggregate process defined in Proposition 1 is given by:*

$$- \text{ If } q > 1, \text{ then } \sum_{k=1}^{\infty} a_k = \frac{p}{p+q-1};$$

¹⁶Note that Gonçalves and Gouriéroux (1994) study extensively the aggregation of individual AR(1) processes in which φ is Beta-distributed (after an homothety) and provide a discussion regarding aggregate long-memory processes. However they do not consider the sum of the autoregressive coefficients.

- If $q \leq 1$, then $\sum_{k=1}^{\infty} a_k = 1$.

Proposition 4 can be shown as follows. Taking the integral form of $m(r)$ we have

$$1 + m(r) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \int_0^1 \frac{x^{p-1}(1-x)^{q-1}}{1-rx} dx,$$

and thus (by the monotone convergence theorem)

$$\lim_{r \rightarrow 1^-} (1 + m(r)) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \int_0^1 x^{p-1}(1-x)^{q-2} dx.$$

Consequently, this is a convergent integral if and only if $q > 1$. In this case,

$$\lim_{r \rightarrow 1^-} (1 + m(r)) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \frac{\Gamma(p)\Gamma(q-1)}{\Gamma(p+q-1)} = 1 + \frac{p}{q-1}$$

and

$$\sum_{k=1}^{+\infty} a_k = \lim_{r \rightarrow 1^-} \frac{m(r)}{1 + m(r)} = \frac{p}{p+q-1}.$$

On the other hand, it follows that $\lim_{r \rightarrow 1^-} (1 + m(r)) = +\infty$ if and only if $q \leq 1$, and $\sum_{k=1}^{+\infty} a_k = 1$ (Theorem 1).

Uniform distribution We now assume that $p = q = 1$ such that the random variable φ is uniformly distributed over the interval $[0, 1)$ and non central moments are given by $u_k = \frac{1}{k+1}$.

Proposition 5 *Suppose that φ is uniformly distributed over the interval $[0; 1)$. Given the disaggregate model defined in Eqs. (1)–(2), the autoregressive coefficients of the limit aggregate process defined in Proposition 1 are given by:*

$$a_k = \frac{|I_k|}{k!}$$

where $I_k = \int_0^1 x(x-1) \cdots (x-k+1) dx$ has the same sign as $(-1)^{k-1}$ for $k \geq 1$. Moreover, $\sum_{k=1}^{\infty} a_k = 1$.

Proposition 5 can be shown by using either Theorem 1 or a new lemma provided in Appendix C. Notably, the coefficients of the series $a(z) = m(z)/(1 + m(z))$ can be computed

as follows. First, the generating moment series is

$$1 + m(z) = \sum_{k \geq 0} u_k z^k = \sum_{k \geq 0} \frac{z^k}{k+1} = -\frac{\log(1-z)}{z}$$

where $|z| < 1$ so that

$$a(z) = \frac{m(z)}{1 + m(z)} = 1 + \frac{z}{\log(1-z)}.$$

Second the expression of $\frac{z}{\log(1-z)}$ is derived by a power series development of the function a . Indeed, using the notation $\psi(z) = \log(1-z)$, one has

$$\int_0^1 e^{x\psi(z)} dx = \frac{e^{\psi(z)} - 1}{\psi(z)} = -\frac{z}{\log(1-z)}$$

where the power series development of $e^{x\psi(z)} = (1-z)^x$ is defined to be:

$$(1-z)^x = \sum_{n \geq 0} \binom{x}{n} (-1)^n z^n = 1 + \sum_{n \geq 1} \frac{(-1)^n}{n!} z^n x(x-1) \cdots (x-n+1).$$

Since this series converges absolutely for $|z| < 1$ and uniformly for x in $[0, 1]$, one obtains

$$\int_0^1 e^{x\psi(z)} dx = \int_0^1 (1-z)^x dx = 1 + \sum_{n \geq 1} \frac{(-1)^n}{n!} z^n I_n$$

where $I_n = \int_0^1 x(x-1) \cdots (x-n+1) dx$ for $n \geq 1$. Finally,

$$\int_0^1 e^{x\psi(z)} dx = -\frac{z}{\log(1-z)} = 1 - \sum_{k \geq 1} \frac{1}{k!} z^k |I_k|$$

and thus

$$a(z) = \sum_{k \geq 1} \frac{|I_k|}{k!} z^k.$$

Several points are worth commenting. On the one hand, the non-negativeness of the autoregressive coefficients for the uniform distribution allows for the use of a standard Hardy-Littlewood Tauberian result and thus might not require Theorem 1. However the non-negativeness of the autoregressive coefficients is not proved for all probability distributions μ with support $[0, 1)$. On the other hand, using Theorem 1, since the

moving average coefficients of the (limit) aggregate process constitute a harmonic series, the corresponding series diverges and thus the sum of the autoregressive coefficients (of the limit aggregate process) equals one. Finally, Proposition 5 extends the result of Linden (1999) in which the behavior of the aggregate process is studied with the autocorrelation function.

Random polynomial density We consider as a last example the case of the polynomial aggregated AR(1) model. More specifically, we suppose that φ has a continuous distribution over $[0, 1]$ that can be represented by a polynomial of degree $d \geq 1$:

$$f(\varphi) = \sum_{s=0}^d c_s \varphi^s \mathbf{1}_{[0,1]}(\varphi)$$

where $\sum_{s=0}^d \frac{c_s}{s+1} = 1$ (the density has to be integrated to one) and f is non-negative in $[0, 1]$. One key feature of the polynomial density function relative to the Beta distribution is that it can be multi-modal (for polynomial density of third order or above). Similarly to the Beta distribution, two cases are considered in Proposition 6: the sum of the autoregressive coefficients does depend on whether $f(1) = 0$ or $f(1) > 0$ with $f(1) = c_0 + c_1 + \dots + c_d$ is the value of the density at $x = 1$.

Proposition 6 *Suppose that φ has a polynomial density function of order d . Given the disaggregate model defined in Eqs. (1)–(2), the series of the autoregressive coefficients of the limit aggregate process defined in Proposition 1 is given by:*

- If $f(1) = 0$, then

$$a(1) = \sum_{k=1}^{\infty} a_k = 1 - \frac{1}{\sum_{n=0}^{d-1} \sum_{k=0}^n \frac{c_{n-k}}{n+1}};$$

- If $f(1) > 0$, then $\sum_{k=1}^{\infty} a_k = 1$.

Indeed, starting from the polynomial density function of order d , the non-central moments $u_k = \mathbb{E} [\varphi^k]$, $k \geq 0$, are given by:

$$\mathbb{E} [\varphi^k] = \sum_{s=0}^d \frac{c_s}{s+k+1},$$

and thus the generating moment series,

$$1 + m(z) = \sum_{k \geq 0} u_k z^k = \sum_{k=0}^{\infty} \left(\sum_{s=0}^d \frac{c_s}{s+k+1} \right) z^k,$$

is convergent at least for $|z| < 1$. Therefore, one needs to study the convergence of the series

$$1 + m(1) = \sum_{k=0}^{\infty} \sum_{s=0}^d \frac{c_s}{s+k+1} = \lim_{K \rightarrow \infty} \sum_{k=0}^K \sum_{s=0}^d \frac{c_s}{s+k+1}$$

The terms of the double sum $S(K) = \sum_{k=0}^K \sum_{s=0}^d \frac{c_s}{s+k+1}$ form an array with $K+1$ rows and $d+1$ columns. Without loss of generality, we can suppose $K > d$. Let $n = s+k$, we have $n-d \leq k \leq n$. Taking diagonal sums along the lines $s+k = 0, 1, \dots, d+K$ of the array, one can write $S(K) = S_1(K) + S_2(K) + S_3(K)$ with

$$\begin{aligned} S_1(K) &= \sum_{n=0}^d \sum_{k=0}^n \frac{c_{n-k}}{n+1} = c_0 + \frac{c_0 + c_1}{2} + \dots + \frac{c_0 + c_1 + \dots + c_d}{d+1} \\ S_2(K) &= \sum_{n=d+1}^K \sum_{k=n-d}^n \frac{c_{n-k}}{n+1} = \sum_{n=d+1}^K \frac{1}{n+1} \sum_{k=n-d}^n c_{n-k} = f(1) \sum_{n=d+1}^K \frac{1}{n+1} \end{aligned}$$

and

$$\begin{aligned} S_3(K) &= \sum_{n=K+1}^{d+K} \sum_{k=n-d}^n \frac{c_{n-k}}{n+1} = \sum_{n=K+1}^{d+K} \frac{1}{n+1} \sum_{k=n-d}^n c_{n-k} \\ &= \frac{c_1 + c_2 + \dots + c_d}{K+2} + \frac{c_2 + \dots + c_d}{K+3} + \dots + \frac{c_d}{K+D+1}. \end{aligned}$$

The first sum, $S_1(K)$, is finite and independent of K . The second sum, $S_2(K)$, is clearly convergent as $K \rightarrow \infty$ if and only if $f(1) = 0$. Finally, the third sum, $S_3(K)$, is finite and its limit is zero when K approaches infinity. In this respect, the generating moment

series $1 + m(1)$ converges if and only if $f(1) = 0$. In this case its limit is $S_1 = S_1(K)$ with

$$S_1(K) = c_0 + \frac{c_0 + c_1}{2} + \dots + \frac{c_0 + c_1 + \dots + c_d}{d + 1}$$

In contrast, when $f(1) > 0$ the series $\sum_{k \geq 0} u_k$ diverges and thus the sum of the autoregressive coefficients equals one. Consequently, the (limit) aggregate process displays long-range dependence.

In this respect, Proposition 6 provides formally the sufficient condition discussed by Chong (2006), i.e. $f(1) > 0$ is sufficient to establish the long memory properties of the (limit) aggregate process. Moreover, if we assume that the polynomial density function is of order zero and $c_0 = 1$, we end up with the uniform distribution (with $f(1) > 0$) and thus Proposition 5. On the other hand, if we assume that the polynomial density function is of the form $ax(1 - x)$ for $a \neq 0$ and $x \in [0, 1)$, then $f(1) = 0$ and the long-run persistence is given by $a(1)$. In contrast, if the probability density function has a support on $[0; 1]$, the aggregation process leads to a generalized integrated process (Lin, 1991; Granger and Ding, 1996). As a final remark, it is worth emphasizing that any distribution such that the generating function of the sequence of the moments is not convergent leads to a long-memory process characterized by Theorem 2.

5 Conclusion

In this paper, we study the aggregation of individual random AR(1) processes under the assumptions that some common factors exist, the (positive) random autoregressive coefficient takes values in $[0, 1)$ with probability distribution μ , and all of the noncentral moments exist. Notably we show by making use of complex analysis that sum of the autoregressive coefficients equals one in the presence of limit aggregate long memory processes: the divergence of the series of noncentral moments is fully equivalent to the presence of a unit root in the autoregressive representation of the stochastic process defined by the limit of the aggregate process. We then illustrate our results using some prominent examples of distribution for the aggregation of random autoregressive AR(1)

processes. This provides some new insights that might deserve some empirical applications and some theoretical developments, as for instance the disaggregation problem.

Appendix: Proofs

Appendix A

Proof of Proposition 2: Taking that the sequence of the moments

$$u_k = \mathbb{E}(\varphi^k) = \int_{[0,1)} x^k d\mu(x)$$

is positive and decreasing, one has $\lim_{k \rightarrow +\infty} u_k = 0$ (by monotone convergence applied to the sequence $(x^k)_{k \geq 1}$ and thus the radius of convergence of the series $\sum_{k \geq 1} u_k z^k$ is at least 1.

Let (r_n) an increasing sequence in $[0, 1)$ such that $\lim_{n \rightarrow +\infty} r_n = 1$. For $x \in [0, 1)$ the functions $f_n(x) = \frac{r_n x}{1 - r_n x}$ are positive therefore by Fatou's Lemma we get

$$\liminf_n m(r_n) = \liminf_n \int_{[0,1)} f_n d\mu(x) \geq \int_{[0,1)} \liminf_n f_n d\mu(x) = \int_{[0,1)} \frac{x}{1-x} d\mu(x) = \sum_{k=0}^{+\infty} u_k$$

Thus if $\sum u_k$ is divergent then $m(r) \rightarrow +\infty$ when $r \rightarrow 1^-$. Moreover by the Abel Theorem (Titchmarsh, 1939, pp.9-10) if $\sum u_k$ is convergent then $\lim_{r \rightarrow 1^-} m(r)$ is finite and $\lim_{r \rightarrow 1^-} m(r) = \sum u_k$. The equivalence then follows.

Appendix B

In Appendix B, we provide the proof of Theorem 1. In so doing, we proceed with three lemmas. Notably we first define m as an analytic function of the complex variable z in the disc $D(0, 1)$ and rewrite $m(z)$ as an integral on $[0, 1)$ with parameter $z \in D(0, 1)$ (Lemma 1). We then show that it can be continuously extended to $\overline{D(0, 1)} \setminus \{1\}$. Then Lemma 2 proves that the function $a(z) = \frac{m(z)}{1+m(z)}$ is well defined in the disc $D(0, 1)$ (i.e., the denominator does not vanish) so that it can be extended to the whole closed disc $\overline{D(0, 1)}$. In this respect, two cases must be studied, according to the nature of the series $\sum u_k$: if $\sum u_k < \infty$ the function m is continuous in the closed disc, so is a ; and if $\sum u_k = \infty$ then $|m(z)| \rightarrow \infty$ when $z \rightarrow 1$ in $D(0, 1)$, therefore $a(z) \rightarrow 1$ when $z \rightarrow 1$ in

$D(0, 1)$. Finally, Lemma 3 proves that the analytic function a is shown to be univalent and provides the use of a Tauberian condition. Then Theorem 1 is proven.

Let m the function defined in the open disc $D(0, 1) = \{z \in \mathbb{C} \text{ with } |z| < 1\}$ by

$$m(z) = \sum_{n=1}^{+\infty} u_n z^n$$

This function is analytic. Moreover, by a classical theorem, on the boundary $C(0, 1) = \{z \in \mathbb{C} \text{ with } |z| = 1\}$ the series $\sum_{n \geq 1} u_n e^{int}$ is convergent for all $t \in]0, 2\pi[$ and by Abel's theorem :

$$m(e^{it}) = \sum_{n=1}^{+\infty} u_n e^{int} = \lim_{r \rightarrow 1^-} \sum_{n=1}^{+\infty} u_n r^n e^{int}$$

Thus the function m is defined in $\overline{D(0, 1)} \setminus \{1\} = \{z \in \mathbb{C} \text{ with } |z| \leq 1, z \neq 1\}$. By positivity we have

$$\sum_{n=1}^{+\infty} E(\varphi^n) r^n = E\left(\sum_{n=1}^{+\infty} \varphi^n r^n\right)$$

for all $0 \leq r < 1$. Thus for $r \in [0, 1)$ we get

$$m(r) = \sum_{n=1}^{+\infty} u_n r^n = E\left(\frac{\varphi r}{1 - \varphi r}\right) = \int_{[0,1)} \frac{rx}{1 - rx} d\mu(x)$$

Lemma 1 For $z \in \overline{D(0, 1)} \setminus \{1\}$

$$m(z) = \int_{[0,1)} \frac{zx}{1 - zx} d\mu(x)$$

Proof: First we prove that the function $z \mapsto \int_{[0,1)} \frac{zx}{1 - zx} d\mu(x)$ is analytic in the open disc $D(0, 1)$, continuous in $\overline{D(0, 1)} \setminus \{1\}$ and thus is defined.

Let K denote a compact set in $\overline{D(0, 1)} \setminus \{1\}$. This compact is included in $\overline{D(0, 1)} \setminus D(1, \varepsilon)$

with $\varepsilon > 0$ so if $z \in K$ we get $zx \in \overline{D(0, 1)} \setminus D(1, \varepsilon)$ for all $x \in [0, 1)$, and thus

$$|1 - zx| \geq \varepsilon.$$

Therefore for all $z \in K$ and $x \in]0, 1[$

$$\left| \frac{zx}{1 - zx} \right| \leq \frac{1}{\varepsilon} x.$$

This proves the continuity of the function $z \mapsto \int_{[0,1)} \frac{zx}{1 - zx} d\mu(x)$ over K . Moreover, by analyticity of $z \mapsto \frac{zx}{1 - zx}$ for all $x \in [0, 1)$ and the previous boundness condition, the function $z \mapsto \int_{[0,1)} \frac{zx}{1 - zx} d\mu(x)$ is analytic in $D(0, 1)$.

Finally we note that this function coincides with $m(r)$, $r \in [0, 1[$, thus by analytic continuation, we obtain

$$m(z) = \int_{[0,1)} \frac{zx}{1 - zx} d\mu(x)$$

for all $z \in D(0, 1)$, and for $z = e^{it}$ with $t \neq 0$

$$m(e^{it}) = \lim_{r \rightarrow 1^-} \sum_{n=1}^{+\infty} u_n r^n e^{int} = \lim_{r \rightarrow 1^-} \int_{[0,1)} \frac{r e^{it} x}{1 - r e^{it} x} d\mu(x) = \int_{[0,1)} \frac{e^{it} x}{1 - e^{it} x} d\mu(x)$$

□

Extension by continuity of the function $\frac{m}{1+m}$

Lemma 2 *The function $\frac{m}{1+m}$ can be extended to a continuous function over $\overline{D(0, 1)}$.*

Proof: The function $1 + m$ doesn't vanish in $\overline{D(0, 1)}$ because

$$1 + m(z) = \int_{[0,1)} \frac{1}{1 - zx} d\mu(x)$$

and if $z = a + ib$ with $a^2 + b^2 \leq 1$ then $ax < 1$ for all $x \in [0, 1)$ so

$$\operatorname{Re}(1 + m(z)) = \int_{[0,1)} \frac{1 - ax}{(1 - ax)^2 + b^2 x^2} d\mu(x) > 0$$

The function $\frac{m}{1+m}$ is therefore defined and continuous in $\overline{D(0, 1)} \setminus \{1\}$ and analytic in $D(0, 1)$.

In order to study the continuity at the point $z = 1$, we need to consider two cases :

1. If $\sum u_k$ is convergent then the series $\sum u_k z^k$ is normally convergent in $\overline{D(0,1)}$ and the function m is continuous in $\overline{D(0,1)}$ and consequently so is $\frac{m}{1+m}$.
2. If $\sum u_k$ is divergent it is sufficient to prove that $|m(z)| \rightarrow +\infty$ when $z \rightarrow 1$. Indeed

$$\frac{m(z)}{1+m(z)} = \frac{1}{1+\frac{1}{m(z)}} \rightarrow 1$$

and we extend $\frac{m}{1+m}$ by 1 at the point $z = 1$.

We now show that $|m(z)| \rightarrow +\infty$ when $z \rightarrow 1$. In so doing, consider a sequence of points $z_k = a_k + ib_k$ with $a_k \rightarrow 1, b_k \rightarrow 0$ and $a_k^2 + b_k^2 \leq 1$. Then

$$\operatorname{Re}(m(z_k)) = -1 + \int_{[0,1)} \frac{1 - a_k x}{(1 - a_k x)^2 + b_k^2 x^2} d\mu(x).$$

As $a_k x \leq 1$ for all integer k and all $x \in [0, 1)$ the functions

$$f_k : x \mapsto \frac{1 - a_k x}{(1 - a_k x)^2 + b_k^2 x^2}$$

are positive in $[0, 1)$. Using Fatou's Lemma, we get

$$\liminf_k \int_{[0,1)} f_k d\mu(x) \geq \int_{[0,1)} \liminf_k f_k d\mu(x)$$

where

$$\liminf_k f_k(x) = \lim f_k(x) = \frac{1}{1-x}.$$

It is therefore sufficient to remark that if $\sum u_k$ is divergent then

$$\int_{[0,1)} \frac{1}{1-x} d\mu(x) \geq \int_{[0,1)} \frac{x}{1-x} d\mu(x) = \sum_{n=0}^{+\infty} u_n = +\infty.$$

So for all sequence (z_k) converging to 1 in the disc, we have $\operatorname{Re}(m(z_k)) \rightarrow +\infty$ and thus $|m(z_k)| \rightarrow +\infty$. \square

Convergence of an analytic function $D(0, 1)$, which is continuous on $\overline{D(0, 1)}$

Lemma 3 provides the use of a Tauberian condition, which turns to be crucial for the proof of Theorem 1.

Lemma 3 *Let $f(z) = \sum_{k \geq 1} b_k z^k$ an analytic function on $D(0, 1)$, continuous on $\overline{D(0, 1)}$. If f is injective on $D(0, 1)$ then the series $\sum_{k \geq 1} b_k$ is convergent.*

Proof: We proceed in two steps. First, we prove that

$$\frac{1}{n} \sum_{k=1}^n k |b_k| \rightarrow 0.$$

On the one hand, the function f is analytic on $D(0, 1)$ thus the image $U = f(D(0, 1))$ is open in \mathbb{C} and U is included in the compact set $f(\overline{D(0, 1)})$ since f is continuous on $\overline{D(0, 1)}$.

On the other hand, the function f being injective on $D(0, 1)$ we have $f'(z) \neq 0$ for all $z \in D(0, 1)$ thus f is a C^1 diffeomorphism between $D(0, 1)$ and U . By the change of variables formula we get

$$\lambda(U) = \int_{D(0,1)} |f'(x + iy)|^2 dx dy$$

and with the use of polar coordinates we get the finite Lebesgue measure of U as the sum of the series

$$\lambda(U) = \pi \sum_{n=1}^{+\infty} n |b_n|^2$$

The convergence of this last series now implies that $\frac{1}{n} \sum_{k=1}^n k |b_k| \rightarrow 0$.

Finally, to verify this assertion let $N \geq 1$ and for $n > N$ write

$$\frac{1}{n} \sum_{k=1}^n k |b_k| = \frac{1}{n} \sum_{k=1}^N k |b_k| + \frac{1}{n} \sum_{k=N+1}^n k |b_k|.$$

Then by Cauchy-Schwarz inequality we have

$$\begin{aligned} \frac{1}{n} \sum_{k=1}^n k |b_k| &\leq \frac{1}{n} \sum_{k=1}^N k |b_k| + \frac{1}{n} \left(\sum_{k=N+1}^n k \right)^{1/2} \left(\sum_{k=N+1}^n k |b_k|^2 \right)^{1/2} \\ &\leq \frac{1}{n} \sum_{k=1}^N k |b_k| + \frac{1}{n} \left(\sum_{k=N+1}^n k |b_k|^2 \right)^{1/2} \end{aligned}$$

and thus

$$\limsup_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^n k |b_k| \leq \frac{1}{n} \left(\sum_{k=N+1}^n k |b_k|^2 \right)^{1/2}.$$

Since $\frac{1}{n} \left(\sum_{k=N+1}^n k |b_k|^2 \right)^{1/2} \rightarrow 0$ it follows that $\frac{1}{n} \sum_{k=1}^n k |b_k| \rightarrow 0$.

Taking this intermediate result, we are now in a position to prove **Lemma 3**. Indeed let $t_n = \frac{1}{n} \sum_{k=1}^n k b_k$ (with $t_0 = 0$). We have $b_n = (t_n - t_{n-1}) + \frac{1}{n} t_{n-1}$. The series $\sum_{n \geq 1} (t_n - t_{n-1})$ is convergent because

$$\sum_{n=1}^N (t_n - t_{n-1}) = t_N \rightarrow 0,$$

and thus the series $\sum_{n \geq 1} (t_n - t_{n-1})$ is Abel summable. Since the series $\sum_{n \geq 1} b_n$ is also Abel summable by continuity of f , we get the Abel-summability of the series $\sum_{n \geq 1} \frac{1}{n} t_{n-1}$.

Finally, since

$$n \left(\frac{1}{n} t_{n-1} \right) = t_{n-1} \rightarrow 0,$$

it follows from the classical Tauber's theorem that the series $\sum_{n \geq 1} \frac{1}{n} t_{n-1}$ (and thus $\sum_{n \geq 1} b_n$) is convergent. \square

Proof of Theorem 1

Using **Lemma 2**, the function $f = a = \frac{m}{1+m}$ is defined and continuous on $\overline{D(0,1)}$ and analytic in $D(0,1)$. It follows from **Lemma 3** that it remains to prove that f is injective in $D(0,1)$. Taking

$$\frac{m(z_1)}{1+m(z_1)} = \frac{m(z_2)}{1+m(z_2)} \Leftrightarrow m(z_1) = m(z_2),$$

it is sufficient to prove the injectivity of the function m on $D(0, 1)$. Moreover it is sufficient to prove this injectivity on $D(0, r)$ for all $0 \leq r < 1$: if m is not injective on $D(0, 1)$ then there exists $z_1 \neq z_2$ such that $m(z_1) = m(z_2)$; so m is not injective on $D(0, r)$ where $r > \max(|z_1|, |z_2|)$.

For all $0 \leq r < 1$ the function m is analytic on $D(0, r)$ and continuous on $\overline{D(0, r)}$. Using Darboux's theorem (Burckel, 1979, p. 310), we could then establish the injectivity of m on $D(0, r)$ by showing that m is injective on the circle of radius r with center 0. Indeed, let

$$\varphi(t) = \operatorname{Re}(m(re^{it})) = \int_{[0,1]} \frac{rx \cos(t) - r^2 x^2}{1 - 2rx \cos(t) + r^2 x^2} d\mu(x).$$

We have

$$\varphi'(t) = \sin t \int_{[0,1]} \frac{-rx(1 - r^2 x^2)}{(1 - 2xr \cos t + r^2 x^2)^2} d\mu(x).$$

We see that the function φ is decreasing on $]0, \pi[$ et increasing on $]\pi, 2\pi[$. It is symmetric across π : we have $\varphi(t) = \varphi(2\pi - t)$. Therefore the only points $t_1 \neq t_2$ with $\varphi(t_1) = \varphi(t_2)$ are the pairs $(t, 2\pi - t)$ with $t \in [0, \pi[$. Moreover we have

$$\operatorname{Im}(m(re^{it})) = \int_{[0,1]} \frac{xr \sin(t)}{1 - 2xr \cos(t) + r^2 x^2} d\mu(x)$$

hence

$$\operatorname{Im}(m(re^{it})) = -\operatorname{Im}(m(re^{i(2\pi-t)}))$$

Since we have

$$\sin(t) \int_{[0,1]} \frac{xr}{1 - 2xr \cos(t) + r^2 x^2} d\mu(x) > 0$$

for $t \in [0, \pi)$, we can't have

$$\operatorname{Im}(m(re^{it})) = \operatorname{Im}(m(re^{i(2\pi-t)}))$$

Therefore $t \mapsto m(re^{it})$ is injective. By the Abel Theorem the sum of the series $\sum a_n$ is equal to the limit $\lim_{r \rightarrow 1} a(r)$ where $r \in [0, 1)$ \square

References

- [1] Aldous, 1980, *Sminaire de Probabilités, Ecole d'été de Saint-Flour*, Springer Verlag.
- [2] Altissimo, F., Mojon, B., Zaffaroni, P., 2009. Can Aggregation Explain the Persistence of Inflation? *Journal of Monetary Economics* 56, 231–241.
- [3] Beran, J., 1994, *Statistics for long-memory processes*, Monograph on statistics and applied probability 61, Chapman and Hall.
- [4] Beran, J., Feng, Y., Ghosh, S. and R. Kulik, 2013. *Long-Memory Processes—Probabilistic Properties and Statistical Methods*, Springer Verlag.
- [5] Brockwell, P.J., and R.A. Davis, 2002. *Introduction to Time Series and Forecasting*, second edition, Springer-Verlag, New York.
- [6] Burckel, R.B., 1979. *An Introduction to Classical Complex Analysis, Vol. I*, Birkhäuser.
- [7] Chamberlain, G. and M. Rothschild, 1983. Arbitrage, factor structure and mean-variance analysis in large asset market. *Econometrica* 51, 1305–1324.
- [8] Chong, T.T.L., 2006. The Polynomial Aggregated AR(1) Model, *Econometrics Journal*, 9, 98–122.
- [9] Chudik, A., Pesaran, M.H., and E. Tosetti, 2011. Weak and Strong Cross-section Dependence and Estimation of Large Panels, *Econometrics Journal*, vol. 14(1), C45-C90.
- [10] Faust, J., and E.M. Leeper, 1997. When Do Long-Run Identifying Restrictions Give Reliable Results?, *Journal of Business & Economic Statistics*, vol. 15, 345–353.
- [11] Forni, M., Lippi, M., 1997. *Aggregation and the Microfoundations of Dynamic Macroeconomics*, Clarendon Press, Oxford.
- [12] Gabaix, X., 2011. The Granular Origins of Aggregate Fluctuations, *Econometrica* 79, 733–772.

- [13] Gonçalves, E., Gouriéroux, C., 1988. Agrégation de Processus Autorégressifs d'ordre 1, *Annales d'Économie et de Statistique* 12, 127–149.
- [14] Granger, C.W.J., 1980. Long memory Relationships and the Aggregation of Dynamics Models, *Journal of Econometrics*, vol. 14, 227–238.
- [15] Granger, C.W.J., and Z. Ding, 1996. Varieties of long memory models, *Journal of Econometrics*, vol. 73, 61–77
- [16] Granger, C.W.J., Morris, M.J., 1976. Time Series Modelling and Interpretation. *Journal of Royal Statistical Society* 139, 246–257.
- [17] Hardy, G.H., 1949. *Divergent Series, 2nd ed.*, Oxford University Press.
- [18] Imbs, J., Jondeau E., Pelgrin, F., 2011. Sectoral Phillips Curves and the Aggregate Phillips Curve, *Journal of Monetary Economics* 58, 328–344.
- [19] Imbs, J., Muntaz, A., Ravn, M.O., Rey, H., 2005. PPP Strikes Back: Aggregation and the Real Exchange Rate, *Quarterly Journal of Economics* 120, 1–43.
- [20] Jondeau, E., Pelgrin, F., 2014a. Aggregating rational expectations models in the presence of heterogeneity, mimeo.
- [21] Jondeau, E., Pelgrin, F., 2014b. Estimating Aggregate Autoregressive Processes When Only Macro Data Are Available, *Economics Letters*, vol. 124, 341–347.
- [22] Korevaar, J., 2004. *Tauberian Theory*, Springer.
- [23] Leipus, R., Philippe, A., Puplinskaitė, D. and D. Surgailis, 2013. Aggregation and Long Memory: Recent Developments, Working Paper, University of Nantes.
- [24] Lewbel, A., 1994. Aggregation and Simple Dynamics, *American Economic Review* 84, 905–918.
- [25] Lin, J-L., 1991. Generalized Integrated Processes and the Aggregation of Dynamic Time Series, *Academia Economic Papers*, vol. 19, 207–226.
- [26] Linden, M., 1999. Time Series Properties of Aggregated AR(1) Processes with Uniformly Distributed Coefficients, *Economics Letters* 64, 31–36.

- [27] Pesaran, M.H., 2003. Aggregation of Linear Dynamic Models: An Application to Life-Cycle Consumption Models under Habit Formation, *Economic Modelling* 20, 383–415.
- [28] Pesaran, M.H., Chudik, A., 2014. Aggregation in Large Dynamic Panels, *Journal of Econometrics* 178, 273–285.
- [29] Pesaran, M. H., and E. Tosetti, 2009. Large panels with common factors and spatial correlation, mimeo, University of Cambridge.
- [30] Pötscher, B.M., 2002. Lower Risk Bounds and Properties of Confidence Sets for Ill-Posed Estimation Problems with Applications to Spectral Density and Persistence Estimation, Unit Roots, and Estimation of Long Memory Parameters, *Econometrica*, vol. 70, 1035–1065.
- [31] Robinson, P.M., 1978. Statistical Inference for a Random Coefficient Autoregressive Model, *Scandinavian Journal of Statistics* 5, 163–168.
- [32] Sims, C., 1971. Distributed Lag Estimation When the parameter Space is Explicitly Infinite-Dimensional, *Annals of Mathematical Statistics*, vol. 42, 1622–1636.
- [33] Sims, C., 1972. The Role of Approximation Prior Restrictions in Distributed Lag Estimation, *Journal of the American Statistical Association*, vol. 67, 169–175.
- [34] Titchmarsh, E.C., 1939. *The Theory of Functions*, 2nd ed, Oxford University Press.
- [35] Zaffaroni, P., 2004. Contemporaneous Aggregation of Linear Dynamic Models in Large Economies, *Journal of Econometrics* 120, 75–102.