Entropy and complexity of compartmental systems

Holger Metzler · Carlos A. Sierra

Abstract Linear compartmental models are commonly used in different areas of science, particularly in modeling the cycles of carbon and other biogeochemical elements. The representation of these models as linear autonomous compartmental systems is useful for comparisons of different model structures and parameterizations on a macroscopic scale. The interpretation of such models as continuous-time Markov chains allows a deeper model analysis on a microscopic scale. In particular we can asses the uncertainty of a single particle's path as it travels through the system as described by path entropy and entropy rate. Path entropy measures the uncertainty of the entire path of a traveling path from its entry into the system until its exit, whereas entropy rate measures the average uncertainty of the instantaneous future of a particle while it is in the system. We derive explicit formulas for these two types of entropy for compartmental systems in equilibrium based on Shannon information entropy and show how they can be used to assess the complexity of such models. Model complexity based on entropy can in turn be used to resolve the problems of equifinality and structural identification in the realm of model selection by means of the maximum entropy principle. We derive the entropy formulas by three/two different approaches, each one allowing different views on and insights into mass-balanced systems.

Holger Metzler
Department of Crop Production Ecology, Swedish University of Agricultural Sciences, Ulls väg 16, 75651 Uppsala, Sweden, E-mail: holger.metzler@slu.de

Carlos A. Sierra
Max Planck Institute for Biogeochemistry, Hans-Knöll-Str. 10, 07745 Jena, Germany, Tel.: +49-3641-576133, E-mail: csierra@bgc-jena.mpg.de
Department of Ecology, Swedish University of Agricultural Sciences, Ulls väg 16, 75651 Uppsala, Sweden, E-mail: carlos.sierra@slu.se

Mathematics Subject Classification (2000) 34A30 · 60J28 · 60K20 · 92B05

1 Introduction

Compartmental models are widely used to describe a range of biological and physical processes that rely on mass conservation principles (**??**). In particular, most models that describe the carbon cycle can be generalized as compartmental models (**???**). These models are very important for predicting interactions between ecosystems and the global climate; however, the predictions of these models diverge widely (**??**) since models have different structures and parameter values. To compare diverse models, we can consider key quantities such as system age and transit time. These quantities provide relevant information about the time scales at which material is cycled in compartmental systems, and facilitate comparisons among different model structures and parameterizations. Formulas for their means were recently provided by **?** for linear nonautonomous compartmental systems. For autonomous systems, that is, systems with constant coefficients, the existence of explicit solutions raises hope for explicit formulas not only for the means, but also for the densities of system age and transit time.

A first attempt in this direction was started by **?** who established formulas for transit time and age densities in dependence on a not explicitly known system response function. This response function was the basis for **?** to compute the desired densities numerically by long-term simulations in two carbon-cycle models. Impulsive inputs and how systems respond to them were later investigated by **?** to obtain explicitly the system response function for a set of carbon-cycle models of very simple structure. Despite the simple structure of these systems, the derivation of the according density formulas is involved because the system's response needs to be transformed from the phase domain to the Laplace domain and back.

In this manuscript, we consider the dynamics of linear autonomous compartmental systems as a stochastic process to show that the impulse response function is the probability density function of a phase-type distribution. To this end, we relate open linear autonomous compartmental systems to absorbing continuous-time Markov chains.

This manuscript is organized as follows. In Sect. **??** we introduce linear autonomous compartmental systems and frame the idea of looking at them from the perspective of one single particle. The transit time of this particle through the system coincides with the absorption time of a continuous-time Markov chain. We shortly present the basic ideas of continuous-time Markov chains in Sect. **??**. The probability distribution of the absorption time is as an example of a phase-type distribution and we compute its cumulative distribution function, probability density function, and moments. In Sect. **??** we use concepts from renewal theory and construct a regenerative process which is the basis for the computation of the system and compartment age densities.

The relationship of Markov chains and linear autonomous compartmental systems is established in Sect. **??**. We use our previous probabilistic results to obtain simple explicit formulas for the densities of the system age, compartment age, and transit time. In Sect. **??** we apply the derived formulas to two autonomous terrestrial carbon-cycle models. Both models are considered in steady state. This allows us to treat a linear model (**?**) as well as a nonlinear one (**?**). Appendix **??** contains the proof of the main theorem from Sect. **??**, and Appendix **??** covers some examples for systems with simple structure.

2 Linear autonomous compartmental systems

Compartmental differential equations are useful tools to describe flows of material between units called compartments under the constraint of mass conservation. Following **?**, a *compartment* is an amount of some material that is kinetically homogeneous. By kinetically homogeneous we mean that the material of a compartment is at all times homogeneous; any material entering the compartment is instantaneously mixed with the material already there. Hence compartments are always *well-mixed*.

We consider the d-dimensional inhomogeneous linear system of differential equations

$$\frac{\mathrm{d}}{\mathrm{d}\,t}\,\mathbf{x}(t) = \mathsf{A}\,\mathbf{x}(t) + \mathbf{u}, \quad t > 0,$$
$$\mathbf{x}(0) = \mathbf{x^0} \in \mathbb{R}^d. \tag{1}$$

We want this equation system to model how mass flows at time $t = 0$ into a set of compartments, afterwards mass flows among the compartments, and eventually leaves the system. Therefore, we require the vector $\mathbf{x^0}$ of initial content and the constant input vector $\mathbf{u}$ to have nonnegative entries only. The law of conservation of mass imposes restrictions on the $d \times d$ matrix A.

Definition 1 A matrix $\mathsf{A} \in \mathbb{R}^{d \times d}$ is called a *compartmental matrix* if

   (i) every off-diagonal entry is nonnegative;
   (ii) each diagonal entry is nonpositive;
(iii) all column sums are nonpositive.

If now A is a compartmental matrix, we call (**??**) a *linear autonomous compartmental system*. The term *autonomous* refers to the fact that both A and $\mathbf{u}$ are independent of time t. A detailed treatment of such systems can be found in **?** and **?**.

At time $t \geq 0$ the vector $\mathbf{x}(t) \in \mathbb{R}^d$ represents the content of the different compartments at time t. The off-diagonal entries a_{ij} of A are called *fractional transfer coefficients*. They are the rates at which mass moves from compartment j to compartment i. For $j = 1, \ldots, d$ the nonnegative value $z_j = -\sum_{i=1}^{d} a_{ij}$ is the rate at which mass leaves the system from compartment j. If at least one of these *output* or *release rates* is greater than zero, system (**??**) is called *open*, otherwise it is called *closed*. We consider open systems only, since closed systems with positive input accumulate mass indefinitely.

2.1 The one-particle perspective

If we look at the behavior of the entire system, the initial content $\mathbf{x^0}$ distorts what we see happening to the system content. There are two ways of getting rid of this disturbing influence of the initial content. One way is to consider the system after it has run for an infinite time such that all the initial content has left. Another way is to look at one single particle that just arrives at the system through new input $\mathbf{u}$.

Since all compartments are well-mixed, this particle's way through the system is not influenced by the presence or absence of any other particles. It enters the system at a compartment according to $\mathbf{u}$ and then, at each time step, it decides on basis of its current position and its schedule whether to stay or to move on. If the decision is to move on, then it can move to another compartment or leave the system, depending only on the connections of the current compartment. The schedule and map followed by the particle are given by the matrix A. The diagonal entries tell the particle how long to stay in a certain compartment, and the off-diagonal entries provide the connections to other compartments. By leaving the system, the particle finishes a cycle and starts a new one by reentering the system.

At each cycle, the sequence of compartments to which the particle belongs at successive time steps constitutes a stochastic process called *discrete-time Markov chain*. Letting the size of the time steps tend to zero, the particle has continuously to decide what to do next. The path of the particle traveling through the system is then represented by a *continuous-time Markov chain* (**?**).

In the next section we introduce the structure and properties of continuous-time Markov chains, and we should always have in mind the traveling particle. When the Markov chain changes its state from j to i, the particle is to be considered to move from compartment j to compartment i. When the Markov chain is absorbed, the particle leaves the system.

The time that elapses from the moment the particle enters the system to the moment of its exit is called *transit time*. To get a grasp on it, from the beginning of the cycle we have to look into the future only.

While the particle is still in the system, the time that has passed since the particle's entry is called its *system age*. If we consider the ages of the particle at a random time, we have to look into the past of the particle, not knowing when it entered the system the last time. Consequently, we have to assume that the particle has ran through the system already infinitely often. Otherwise we would imply the existence of a maximum age of the particle, but any choice of this maximum value would be arbitrary and ill-founded. Therefore, we will consider a particular regenerative process: a sequence of continuous-time Markov chains.

₁₄₉ 2.2 From one particle to all particles in the system

₁₅₀ Let us assume that the system at time $t \geq 0$ contains n particles with sys-
₁₅₁ tem ages represented by n independent and identically distributed random
₁₅₂ variables. Hence, the system age A_k of particle k is assumed to have the cu-
₁₅₃ mulative distribution function F_A for all $k = 1, 2, \ldots, n$. The share of particles
₁₅₄ with system age less than or equal to $y \geq 0$ equals

$$F_n(y) = \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{(-\infty, y]}(A_k).$$

₁₅₅ Here $\mathbb{1}_M(m)$ denotes the indicator function. It equals 1 if $m \in M$ and it
₁₅₆ equals 0 otherwise. The function F_n is called *empirical distribution function*
₁₅₇ of $A_1, A_2, \ldots, A_n$, and from the Glivenko-Cantelli theorem (**?**) we know that
₁₅₈ it converges almost surely uniformly to F_A if the number n of particles goes to
₁₅₉ infinity. Consequently, if the system contains an infinite amount of particles,
₁₆₀ the share of the total system content $\mathbf{x}(t)$ that has system age less than or
₁₆₁ equal to y is

$$F_A(y) = \int_0^y f_A(\tau) \, \mathrm{d}\tau,$$

₁₆₂ where f_A is the common probability density function of the particles' ages.
₁₆₃ We can thus write

$$\mathbf{x}(t) = \mathbf{x}(t) \int_0^\infty f_A(y) \, \mathrm{d}y$$

₁₆₄ to represent the distribution of system ages over the system.

3 Continuous-time Markov chains

₁₆₆ Markov chains are the most important examples of random processes, because
₁₆₇ their simple structure and high diversity allow their applications to many
₁₆₈ different problems. In particular, they are the simplest mathematical models
₁₆₉ for random phenomena evolving in time. Their characteristic property is that
₁₇₀ they retain no memory on the states of the system in the past. Only the
₁₇₁ current state of the process can influence where it goes next. If the process
₁₇₂ can assume only a finite or countable set of states, it is called a Markov chain.
₁₇₃ Discrete-time Markov chains are usually defined on a set of integers, whereas
₁₇₄ continuous-time Markov chains live on a subset of the real line. In this section
₁₇₅ we introduce the basic theory of Markov chains along the line of **?**, from which
₁₇₆ also most unproven results are taken.
₁₇₇ We say that a vector $\boldsymbol{\lambda} \in \mathbb{R}^d$ is a *distribution* if all its components are
₁₇₈ nonnegative and sum to one. Furthermore, we call a matrix $\mathsf{P} = (p_{ij})_{i,j \in J}$
₁₇₉ *stochastic* on a finite set J if every column $(p_{ij})_{i \in J}$ is a distribution. Note
₁₈₀ that in standard literature on Markov chains row sums are considered. We

use column sums instead because thereby the connection to compartmental matrices will become more obvious later.

We shall now formalize the rules for a Markov chain by a definition in terms of the corresponding transition rate matrix.

Definition 2 Let J be a finite state-space. A *Q-matrix on J* is a matrix $\mathsf{Q} = (q_{ij})_{i,j \in J}$ satisfying the following conditions:

(i) $q_{ij} \geq 0$ for all $i \neq j$;

(ii) $0 \leq -q_{jj} < \infty$ for all j;

(iii) $\sum\limits_{i \in J} q_{ij} = 0$ for all j.

We interpret each off-diagonal entry q_{ij} as the *rate of going from state j to state i* and the diagonal entries q_{jj} as the *rate of leaving* state j.

Definition 3 Let Q be a Q-matrix and $\boldsymbol{\lambda}$ a distribution on a finite state-space J. We say that a stochastic process $X = (X_t)_{t \geq 0}$ is a *continuous-time Markov chain* on J with *transition rate matrix* Q and *initial distribution* $\boldsymbol{\lambda}$ if the following conditions hold:

(i) $\mathbb{P}(X_0 = j) = \lambda_j$ for all $j \in J$;

(ii) for all $n = 0, 1, 2, \ldots$, all times $0 \leq t_0 \leq t_1 \leq \cdots \leq t_{n+1}$, and $j_0, j_1, \ldots, j_{n+1} \in J$,

$$\mathbb{P}(X_{t_{n+1}} = j_{n+1} \,|\, X_{t_0} = j_0, \ldots, X_{t_n} = j_n) = e^{(t_{n+1} - t_n)\,\mathsf{Q}}_{j_n j_{n+1}},$$

where for any matrix A the expression e^{A} denotes the matrix exponential.

Since Q does not depend on time, X is called *time invariant.* Property (**??**) states that the future evolution of a Markov process depends only on its current state and not on its history. This is called *Markov property.*

Let X be a continuous-time Markov chain on J with transition rate matrix Q and initial distribution $\boldsymbol{\lambda}$. Then, for $i, j \in J$, the probability of being in state i at time t having started in state j is equal to

$$\mathbb{P}(X_t = i \,|\, X_0 = j) = e^{t\,\mathsf{Q}}_{ij}$$

and the unconditional probability of being in state i at time t is

$$\mathbb{P}(X_t = i) = \left(e^{t\,\mathsf{Q}}\,\boldsymbol{\lambda}\right)_i.$$

Sometimes we may write $\mathbb{P}_j(X_t = i)$ for $\mathbb{P}(X_t = i \,|\, X_0 = j)$.

Closely connected to a continuous-time Markov chain X is its embedded jump chain. This is a discrete-time Markov chain $Y = (Y_n)_{n=0,1,2,\ldots}$ that keeps track of the jumps of X.

Definition 4 Let $\boldsymbol{\lambda}$ be a distribution and P a stochastic matrix on a finite state-space J. A stochastic process $Y = (Y_n)_{n=0,1,2,\ldots}$ is called a *discrete-time Markov chain* with state-space J, *initial distribution* $\boldsymbol{\lambda}$, and *transition probability matrix* $\mathsf{P} = (p_{ij})_{i,j \in J}$ if for $n = 0, 1, 2, \ldots$, and $j_0, j_1 \ldots, j_{n+1} \in J$,

(i) $\mathbb{P}(Y_0 = j_0) = \lambda_{j_0}$;

(ii) $\mathbb{P}(Y_{n+1} = j_{n+1} \mid Y_0 = j_0, Y_1 = j_1, \ldots, Y_n = j_n) = p_{j_{n+1}j_n}$.

Note that Y is time invariant, since the transition probability matrix does not depend on n.

The off-diagonal entries p_{ij} are the probabilities of changing from state j to state i in the next step and the diagonal entries p_{jj} are the probabilities of remaining in state j.

The *jump chain* of a continuous time Markov chain is now defined such that $Y_0 = X_0$ and every time t the process X jumps into another state, the process Y takes the new value of X_t. So Y_n represents the state of the process X after the nth jump. Clearly, the state-space of Y is also J. The transition probability matrix $\Pi = (\pi_{ij})_{i,j\in J}$ of the jump chain is given by

$$\pi_{ij} = \begin{cases} -q_{ij}/q_{jj} & \text{if } j \neq i \text{ and } q_{jj} \neq 0, \\ 0 & \text{if } j \neq i \text{ and } q_{jj} = 0, \end{cases}$$

$$\pi_{jj} = \begin{cases} 0 & \text{if } q_{jj} \neq 0, \\ 1 & \text{if } q_{jj} = 0. \end{cases}$$

3.1 Absorbing continuous-time Markov chains

Let us consider a continuous-time Markov chain $X = (X_t)_{t\geq 0}$ with a special structure. Its finite state-space J is supposed to be equal to $\{1, 2, \ldots, d, d+1\}$ for some natural number $d \geq 1$, and its transition rate matrix has the following shape:

$$\mathsf{Q} = \begin{pmatrix} \mathsf{A} & \mathbf{0} \\ \mathbf{z}^T & 0 \end{pmatrix} \in \mathbb{R}^{(d+1)\times(d+1)}. \tag{2}$$

Here, $\mathbf{0}$ is the d-dimensional column vector containing only zeros. Let $S = \{1, 2, \ldots, d\} \subseteq J$. The $d \times d$-matrix $\mathsf{A} = (a_{ij})_{i,j\in S}$ is supposed to meet the requirements (**??**) and (**??**) of a Q-matrix, but instead of property (**??**) of Definition **??** it fulfills only the weaker condition

$$\sum_{i\in S} a_{ij} \leq 0 \text{ for all } j \in S.$$

Since Q is required to be a Q-matrix, the vector $\mathbf{z} \in \mathbb{R}^d$ must contain the missing parts to make the columns sum to zero. Consequently, $z_j = -\sum_{i\in S} a_{ij}$ or, in matrix notation,

$$\mathbf{z}^T = -\mathbf{1}^T \mathsf{A}, \tag{3}$$

where $\mathbf{1}^T$ denotes the d-dimensional row vector consisting of ones. This means that the z_j are nonnegative and denote the transition rates from j to $d+1$. The $(d+1)$st column of Q comprises only zeros. Consequently, the process X cannot change its state anymore once it has reached state $d+1$. For that reason $d+1$ is called the *absorbing state* of X. We exclude the trivial case

in which the process starts in its absorbing state by considering only initial distributions $\boldsymbol{\lambda}$ with $\lambda_{d+1} = 0$ and define

$$\boldsymbol{\beta} := (\lambda_1, \lambda_2, \ldots, \lambda_d)^T \tag{4}$$

to be the new initial distribution of X.

A standard linear algebra argument shows that

$$e^{t\,\mathsf{Q}} = \begin{pmatrix} e^{t\,\mathsf{A}} & \mathbf{0} \\ * & 1 \end{pmatrix}, \quad t \geq 0,$$

where the asterisk $*$ is a place holder for a d-dimensional row vector. This means that, for $i, j \in S$,

$$\mathbb{P}_j(X_t = i) = e_{ij}^{t\,\mathsf{A}}$$

and

$$\mathbb{P}(X_t = i) = \left(e^{t\,\mathsf{A}} \boldsymbol{\beta}\right)_i. \tag{5}$$

It is an interesting question whether the absorbing state will always be reached no matter in which state the process starts. In this case the *absorption time T*, that is the time at which the process changes to the absorbing state, is finite with probability one. From **?** we have the following lemma.

Lemma 1 *If A is nonsingular, then the absorption time T is finite with probability one.*

Proof Let A be nonsingular. Then $a_{jj} < 0$ for all $j \in S$, because if $a_{jj} = 0$ for a $j \in S$, then the entire jth column must be zero and thus A must be singular, which is a contradiction. Consequently, we can assume that all diagonal entries of A are strictly negative. Let us consider the jump chain Y of X. Clearly, also for the jump chain $d + 1$ is the absorbing state.

Now, define the random variable

$$H := \inf\{n : Y_n = d + 1\}$$

as the number of the jump of process X in which the absorbing state is reached. For $j \in S$ let q_j denote the probability of eventual absorption into state $d + 1$, starting in state j. Then $q_j = \mathbb{P}_j(T < \infty) = \mathbb{P}_j(H < \infty)$.

Since Y is a time invariant Markov chain, for $j \in S$, we have

$$\mathbb{P}_j(H < \infty \,|\, Y_1 = i) = \mathbb{P}_i(H < \infty) = \begin{cases} q_i, & i \in S, \\ 1, & i = d+1, \end{cases}$$

and

$$\begin{aligned} \mathbb{P}_j(H < \infty) &= \sum_{i=1}^{d+1} \mathbb{P}_j(H < \infty, Y_1 = i) \\ &= \sum_{i=1}^{d+1} P_j(H < \infty \,|\, Y_1 = i)\,\mathbb{P}_i(Y_1 = i) \\ &= \sum_{\substack{i=1, \\ i \neq j}}^{d} q_i\, \frac{a_{ij}}{-a_{jj}} + \frac{z_j}{-a_{jj}}. \end{aligned}$$

Hence, the eventual absorption probabilities satisfy the system of linear equations

$$q_j = -\sum_{\substack{i=1, \\ i \neq j}}^{d} q_i \frac{a_{ij}}{a_{jj}} - \frac{z_j}{a_{jj}}, \quad j \in S,$$

or, equivalently,

$$\mathsf{A}^T \mathbf{q} + \mathbf{z} = \mathbf{0}.$$

Using $\mathbf{z} = -\mathsf{A}^T \mathbf{1}$, this is equivalent to

$$\mathsf{A}^T (\mathbf{q} - \mathbf{1}) = \mathbf{0},$$

which has the unique solution $\mathbf{q} = \mathbf{1}$ if A is nonsingular. Hence, absorption from any state is certain. $\qquad\square$

Definition 5 A Markov chain X on $J = \{1, 2, \ldots, d, d+1\}$ whose transition rate matrix Q has the structure (**??**) and that will eventually be absorbed to $d+1$ with probability one is called *absorbing*. The matrix A is called its *transition rate matrix*, $S = \{1, 2, \ldots, d\}$ its *state-space*, and $d+1$ its *absorbing state*. If $\boldsymbol{\lambda}$ denotes the initial distribution of X, then $\boldsymbol{\beta}$ as defined in (**??**) is called the *initial distribution* of the absorbing chain.

The following results related to aborbing continuos-time Markov chains will be used in the forthcoming.

Remark 1 We know that for an absorbing continuous-time Markov chain eventual absorption is certain, independent of the initial state. Hence, for all $i, j \in S$,

$$\lim_{t \to \infty} \mathbb{P}_j(X_t = i) = \lim_{t \to \infty} e_{ij}^{t\mathsf{A}} = 0.$$

Lemma 2 *For any nonsingular matrix* A*, we have*

$$\int\limits_{0}^{\infty} e^{t\mathsf{A}} \, \mathrm{d}t = -\mathsf{A}^{-1}. \tag{6}$$

Remark 2 Note that since Q is a Q-matrix, $e^{t\mathsf{Q}}$ is stochastic. This makes all of its entries and hence all entries of $e^{t\mathsf{A}}$ nonnegative. From (**??**) we can see that also all entries of $-\mathsf{A}^{-1}$ are nonnegative.

From now on let A be the nonsingular transition rate matrix of an absorbing continuous-time Markov chain $X = (X_t)_{t \geq 0}$ on $S = \{1, 2, \ldots, d\}$ with initial distribution $\boldsymbol{\beta}$ and absorbing state $d+1$.

282 3.2 Occupation time

283 Before its absorption, X takes on different states $j \in S$. Its *occupation time* of
284 state j is defined as the time the process spends in state j. It is given as the
285 nonnegative random variable

$$O_j := \int_0^\infty \mathbb{1}_{\{X_t = j\}} \, \mathrm{d}t.$$

Using $\mathbb{E}[\mathbb{1}_{X_t = j}] = \mathbb{P}(X_t = j)$ and (**??**), we can compute its expected value to

$$\mathbb{E}[O_j] = \left(-\mathsf{A}^{-1}\,\boldsymbol{\beta}\right)_j.$$

286 3.3 The last state before absorption

287 Let E denote the state from which X jumps to the absorbing state. Owing to
288 the time invariance of X we have

$$\mathbb{P}(E = j) = z_j \, \mathbb{E}[O_j], \tag{7}$$

289 where O_j is the occupation time of state j.

290 3.4 The absorption time

291 The *absorption time* T is the random variable that tells the time, when the
292 absorbing state is reached:

$$T := \inf\{t \geq 0 : X_t = d + 1\}.$$

293 Another name for it is *hitting time* of the absorbing state. We are interested
294 in the probability distribution of T and it turns out that T follows a *phase-*
295 *type distribution* that depends on the transition rate matrix A and the initial
296 distribution $\boldsymbol{\beta}$.

297 *3.4.1 Probability distributions of phase-type*

298 Phase-type distributions constitute a highly versatile class of probability distri-
299 butions and are closely related to the solutions of systems of linear differential
300 equations with constant coefficients. As mixtures of exponential distributions
301 they generalize the Erlang, the hypoexponential, and the hyperexponential
302 distribution. Those are widely used in queuing theory and the field of renewal
303 processes. An introduction to phase-type distributions and the unifying matrix
304 formalism used in this paper can be found in **?**.
305   The *cumulative distribution function* of the absorption time T is defined
306 as

$$F_T(t) = \mathbb{P}(T \leq t), \text{ for } t \geq 0,$$

307 so it describes the probability of the process X to have reached the absorbing
308 state before or at time t. This is equivalent to saying that at time t the process
309 X is not in any of the states $j \in S$, which by means of (**??**) gives

$$F_T(t) = 1 - \sum_{j \in S} \mathbb{P}(X_t = j) = 1 - \mathbf{1}^T e^{t\,\mathsf{A}}\,\boldsymbol{\beta}, \quad t \geq 0.$$

310 Using $\mathbf{z}^T = -\mathbf{1}^T \mathsf{A}$ from (**??**), we get the *probability density function*

$$f_T(t) = \mathbf{z}^T e^{t\,\mathsf{A}}\,\boldsymbol{\beta}, \quad t \geq 0. \tag{8}$$

311 A probability distribution according to this density is called *phase-type*
312 with transition rate matrix A and initial distribution $\boldsymbol{\beta}$. We write $T \sim \mathrm{PH}(\boldsymbol{\beta}, \mathsf{A})$.

313 *3.4.2 Moments of the phase-type distribution*

Let $T \sim \mathrm{PH}(\boldsymbol{\beta}, \mathsf{A})$ with nonsingular A and $\mathbf{z}^T = -\mathbf{1}^T \mathsf{A}$. We can use (**??**) to
calculate the nth moment of the phase-type distribution. Repeated integration
by parts leads to

$$\mathbb{E}\left[T^n\right] = \int_0^\infty t^n\, f_T(t)\,\mathrm{d}t = (-1)^n\, n!\, \mathbf{1}^T \mathsf{A}^{-n}\,\boldsymbol{\beta}.$$

314 In particular, the first moment of the phase-type distribution is

$$\mathbb{E}[T] = -\mathbf{1}^T \mathsf{A}^{-1}\,\boldsymbol{\beta} = \|\mathsf{A}^{-1}\,\boldsymbol{\beta}\|, \tag{9}$$

315 where $\|\mathbf{v}\| := \sum_{j=1}^d |v_j|$ denotes the norm of a vector $\mathbf{v} \in \mathbb{R}^d$. The last equality
316 holds since both the matrix $-\mathsf{A}^{-1}$ and the vector $\boldsymbol{\beta}$ are nonnegative. For $-\mathsf{A}^{-1}$
317 this follows from Remark **??** and for $\boldsymbol{\beta}$ because it is a distribution.
318 Note that the variance, that is the second central moment, of the phase-
319 type distribution is $\sigma_T^2 := \mathbb{E}[T^2] - (\mathbb{E}[T])^2$.

320 *3.4.3 A closure property of the phase-type distributions*

321 The class of phase-type distributions has several closure properties, one of
322 which is given by the following lemma (**?**).

Lemma 3 *If $T \sim \mathrm{PH}(\boldsymbol{\beta}, \mathsf{A})$, then a random variable with cumulative distri-
bution function*

$$\frac{1}{\mathbb{E}[T]} \int_0^t [1 - F_T(\tau)]\, d\tau$$

323 *is phase-type distributed with transition rate matrix A and initial distribution*

$$\boldsymbol{\eta} = \frac{-\mathsf{A}^{-1}\,\boldsymbol{\beta}}{\|\mathsf{A}^{-1}\,\boldsymbol{\beta}\|}.$$

4 A renewal process and a regenerative process

Assume we restart an absorbing continuous-time Markov chain immediately every time it hits the absorbing state. The goal of this section is to determine the distribution of the age of this new process at a random time. Age means here the time that has passed since the last restart. First, we set up the new situation, then we compute the distribution of the age at a random time, and as a last step we compute the age distribution at a random time under the condition of being in a fixed state. We will deal with elements from renewal theory such as renewal and regenerative processes. A comprehensive treatise of this topic can be found in **?**.

4.1 Definition of the renewal and the regenerative process

Let X be an absorbing continuous-time Markov chain on a finite state-space J with transition rate matrix A and initial distribution $\boldsymbol{\beta}$. We now stop the process X at its absorption time T, execute an immediate restart, and repeat this procedure over and over again. This gives an infinite sequence $(X^k)_{k=1,2,\dots}$ of independent and identically distributed cycle processes, where each cycle process behaves like the process X up to absorption. This sequence constitutes a *regenerative process* $Z = (Z_t)_{t\geq 0}$ with $Z_t = X_t^k$ for $t \in [T_{k-1}, T_k)$, where $T_0 := 0$. The process that counts the number of restarts is called a *renewal process*.

Let us further consider another regenerative process $Y = (Y_t)_{t\geq 0}$ with the same cycle lengths as Z, defined by $Y_t = Y_t^k$ for $t \in [T_{k-1}, T_k)$. This process is assumed to have only two possible states, 1 and 0. We say that Y is *on* at time t if $Y_t = 1$, otherwise it is called *off*. This particular type of regenerative process is called *alternating*. Note that X_t^k and Y_t^k are defined to be zero if $t \notin [T_{k-1}, T_k)$.

The alternating process Y together with the following theorem build the main tool for the computation of the age distributions in which we are interested.

Theorem 1 *At a random time $\tau \geq 0$ the probability of the alternating process Y being on equals the ratio of expected on-time during the first cycle to the average cycle length. Consequently,*

$$P(Y_\tau = 1) = \frac{1}{\mathbb{E}[T]} \, \mathbb{E} \int_0^\infty \mathbb{1}_{\{Y_t^1 = 1\}} \, \mathrm{d}t.$$

Despite the intuitive nature of this result, its proof requires quite some technical effort and is postponed to Appendix **??**.

358 ## 4.2 The age of the regenerative process

359 Consider the age process $(A_t)_{t \geq 0}$ defined such that the random variable A_t
360 describes the time that has passed by t since the last restart. It is also called
361 *backward recurrence time* of the renewal process. It can be considered as the
362 age of the regenerative process Z at time t. For $y \geq 0$ and a randomly chosen
363 time τ we are interested in $F_{A_\tau}(y) = \mathbb{P}(A_\tau \leq y)$.
364   To obtain that probability we construct an alternating process Y that is
365 on at time t if and only if the time elapsed since the last restart is less than
366 or equal to y. Obviously, $\mathbb{P}(A_\tau \leq y) = \mathbb{P}(Y_\tau = 1)$. We take into account that
367 $Y_t^1 = 1$ if and only if $t < T_1$ and $t \leq y$, and see that

$$\mathbb{E} \int_0^\infty \mathbb{1}_{\{Y_t^1 = 1\}} \, dt = \mathbb{E} \int_0^y \mathbb{1}_{\{t < T_1\}} \, dt = \int_0^y \mathbb{P}(T > t) \, dt = \int_0^y [1 - F_T(t)] \, dt.$$

368 We write A instead of A_τ and apply Theorem **??** to the alternating process
369 Y. Now we obtain the cumulative distribution function

$$F_A(y) = \frac{1}{\mathbb{E}[T]} \int_0^y [1 - F_T(t)] \, dt, \quad y \geq 0, \tag{10}$$

370 of the age of the regenerative process Z. Hence, this age is again phase-type
371 distributed by Lemma **??**, now with initial probability vector $\boldsymbol{\eta}$.

372 ## 4.3 The age of the regenerative process in a fixed state

373 Fix $j \in S$ and let a_j be the random variable that describes the age of the
374 process Z given that it is in state j. For $y \geq 0$ that means

$$\mathbb{P}(a_j \leq y) = \mathbb{P}(A_\tau \leq y \,|\, Z_\tau = j).$$

375 The conditional probability on the right hand side can be expressed as

$$\mathbb{P}(A_\tau \leq y \,|\, Z_\tau = j) = \frac{\mathbb{P}(A_\tau \leq y, \, Z_\tau = j)}{\mathbb{P}(Z_\tau = j)}. \tag{11}$$

376 We can compute the numerator and the denominator using the alternating
377 processes Y^1 and Y^2, respectively. The process Y^1 is defined to be on at time
378 t if and only if $t \leq y$, $t < T_1$, and $X_t^1 = j$, whereas Y^2 is defined to be on if
379 and only if Z is in state j. We invoke Theorem **??** twice and get

$$\mathbb{P}(A_\tau \leq y, \, Z_\tau = j) = \frac{1}{\mathbb{E}[T]} \int_0^y \mathbb{P}(X_t = j) \, dt. \tag{12}$$

380 and

$$\mathbb{P}(Z_\tau = j) = \frac{\mathbb{E}[O_j]}{\mathbb{E}[T]}. \tag{13}$$

Plugging Eq. (**??**) for the numerator and Eq. (**??**) for the denominator in (**??**), we have

$$F_{a_j}(y) = \frac{1}{\mathbb{E}[O_j]} \int_0^y \mathbb{P}(X_t = j) \, \mathrm{d}t, \quad y \geq 0.$$

This means that the probability density function of the age of Z in state j given by

$$f_{a_j}(y) = \frac{\mathbb{P}(X_y = j)}{\mathbb{E}[O_j]} = \frac{\left(e^{y\mathsf{A}}\boldsymbol{\beta}\right)_j}{\left(-\mathsf{A}^{-1}\boldsymbol{\beta}\right)_j}, \quad y \geq 0. \tag{14}$$

5 Compartmental systems and Markov chains

We come back to the linear autonomous compartmental system (**??**). Consider the $(d+1) \times (d+1)$ matrix

$$\mathsf{Q} = \begin{pmatrix} \mathsf{A} & \mathbf{0} \\ \mathbf{z}^T & 0 \end{pmatrix}.$$

Owing to the properties of A and $\mathbf{z} = (z_j)_{j=1,2,\ldots,d}$, the extended matrix Q meets Definition **??**. Consequently, it is the transition rate matrix of a continuous-time Markov chain $X = (X_t)_{t \geq 0}$ on the state-space $J = \{1, 2, \ldots, d, d+1\}$.

Assume mass $\mathbf{u} \in \mathbb{R}^d$ comes into system (**??**) at time t_0. Since the system is linear, the way how the mass will be distributed through it can be modeled by a system without input and with initial value $\mathbf{u}$:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathbf{x}}(t) &= \mathsf{A}\,\tilde{\mathbf{x}}(t), \quad t > t_0, \\ \tilde{\mathbf{x}}(t_0) &= \mathbf{u}. \end{aligned} \tag{15}$$

Furthermore, the fractional transfer coefficients of this system are time independent, so we can shift the entire system to the left and consider it having started in time $t_0 = 0$. The content of compartment j at time $t \geq 0$ is given by $\tilde{x}_j(t) = \left(e^{t\mathsf{A}}\mathbf{u}\right)_j$.

Let

$$\boldsymbol{\lambda} := \frac{1}{\|\mathbf{u}\|}(u_1, u_2, \ldots, u_d, 0)^T$$

and $\boldsymbol{\beta} := (\lambda_1, \lambda_2, \ldots, \lambda_d)^T$. Then the probability of the continuous-time Markov chain X with initial distribution $\boldsymbol{\beta}$ being in state $j \in S := \{1, 2, \ldots, d\}$ at time t is, by (**??**),

$$\mathbb{P}(X_t = j) = \left(e^{t\mathsf{A}}\boldsymbol{\beta}\right)_j = \frac{\tilde{x}_j(t)}{\|\mathbf{u}\|}.$$

So, $\mathbb{P}(X_t = j)$ is the proportion of the initially present mass of system (**??**) that is in compartment j at time t. Consequently, the continuous-time Markov chain X describes the stochastic travel of a single particle through the compartmental system. When the traveling particle leaves the compartmental system, the process X jumps to the absorbing state $d + 1$.

408 5.1 Global asymptotic stability

409 We assume the compartmental matrix A to be nonsingular. Then the constant
410 vector $\mathbf{x^*} = \mathsf{A}^{-1}\mathbf{u}$ is a *steady state solution* or *equilibrium* of system (**??**),
411 that is, $\mathrm{d}\mathbf{x^*}/\mathrm{d}t = \mathbf{0}$.
412   We know that any solution $\mathbf{x}$ of the linear system (**??**) has the form (**?**)

$$\mathbf{x}(t) = e^{t\mathsf{A}}\mathbf{x^0} + \int_0^t e^{(t-\tau)\mathsf{A}}\mathbf{u}\,\mathrm{d}\tau, \quad t \geq 0.$$

If we let $t \to \infty$, we can use (**??**) and see from Remark **??** that

$$\lim_{t\to\infty}\mathbf{x}(t) = -\mathsf{A}^{-1}\mathbf{u} + \lim_{t\to\infty}\mathsf{A}^{-1}e^{t\mathsf{A}}\mathbf{u} = -\mathsf{A}^{-1}\mathbf{u},$$

413 which means that all solutions of (**??**) converge to $\mathbf{x^*}$, independently of their
414 initial value $\mathbf{x^0}$. This makes the equilibrium $\mathbf{x^*} = -\mathsf{A}^{-1}\mathbf{u}$ *globally asymp-*
415 *totically stable* and by the classic Lyapunov theorem (**?**, Theorem 2.10) all
416 eigenvalues of A have negative real part.
417   It is well known that system (**??**) is globally asymptotically stable, if A is
418 strictly diagonally dominant. In our case, A is not strictly diagonally domi-
419 nant (it is diagonally dominant, though), but it is a nonsingular compartmen-
420 tal matrix. This means that A is the transition rate matrix of an absorbing
421 continuous-time Markov chain. Consequently, for at least one $j \in S$ the rate
422 z_j of leaving the system is strictly greater than zero. This makes (**??**) an open
423 system (**?**).

424 5.2 Steady state compartment content

425 For $j \in S$ the steady state content of compartment j is $x_j^* = -\left(\mathsf{A}^{-1}\mathbf{u}\right)_j$. We
426 see

$$\frac{x_j^*}{\|\mathbf{u}\|} = -\left(\mathsf{A}^{-1}\boldsymbol{\beta}\right)_j = \mathbb{E}[O_j]. \tag{16}$$

427 Consequently, the steady state compartment content x_j^* is proportional to the
428 expected occupation time of state j by the absorbing continuous-time Markov
429 chain X.

430 5.3 Release from the system

431 Let $\mathbf{r} \in \mathbb{R}^d$ be the vector of mass release or output from the system in steady
432 state. For $j \in S$ the component r_j is the release of mass from compartment j. It
433 can be computed as the rate of mass that leaves compartment j at an arbitrary
434 time $t \geq 0$ and does not go to any other compartment $i \in S$. Consequently,
435 $r_j = z_j\,x_j^*$. Probabilistically, we would expect this value to be connected to the
436 probability of the absorbing continuous-time Markov chain X to be absorbed

through state j. From (**??**) the probability of j being the last state before absorption is $\mathbb{P}(E = j) = -z_j \left(\mathsf{A}^{-1} \boldsymbol{\beta} \right)_j$. We use $\boldsymbol{\beta} = \mathbf{u}/\|\mathbf{u}\|$ and $\mathbf{x}^* = -\mathsf{A}^{-1} \mathbf{u}$ to get

$$\mathbb{P}(E = j) = z_j \, x_j^* \, \frac{1}{\|\mathbf{u}\|} = \frac{r_j}{\|\mathbf{u}\|}.$$

This means that the release through compartment j is proportional to the probability of X being absorbed through state j. Since absorption is certain, $\sum_{j \in S} \mathbb{P}(E = j) = 1$, hence $\|\mathbf{r}\| = \sum_{j \in S} r_j = \|\mathbf{u}\|$, reflecting that in steady state total system output equals total system input.

5.4 Age distribution of the system in steady state

Suppose that the total system content in steady state has an unknown age density expressed by

$$\|\mathbf{x}^*\| = \|\mathbf{x}^*\| \int\limits_0^\infty f_A(y) \, dy.$$

We have already considered a particle that travels over and over again through the system and computed its age at a random time by methods from renewal theory in Sect. **??**. The according regenerative process expresses the need for an infinite history. Consequently, it reflects the behavior of the compartmental system in steady state, when all initial mass has left the system.

The age random variable A was shown to be phase-type distributed with transition rate matrix A and initial distribution

$$\boldsymbol{\eta} = \frac{-\mathsf{A}^{-1} \boldsymbol{\beta}}{\|\mathsf{A}^{-1} \boldsymbol{\beta}\|} = \frac{\|\mathbf{u}\|}{\|\mathbf{u}\|} \frac{-\mathsf{A}^{-1} \mathbf{u}}{\|\mathsf{A}^{-1} \mathbf{u}\|} = \frac{\mathbf{x}^*}{\|\mathbf{x}^*\|}.$$

So the density is given by (**??**) with $\boldsymbol{\eta}$ replacing $\boldsymbol{\beta}$. We have

$$f_A(y) = \mathbf{z}^T \, e^{y \, \mathsf{A}} \, \boldsymbol{\eta} = \mathbf{z}^T \, e^{y \, \mathsf{A}} \, \frac{\mathbf{x}^*}{\|\mathbf{x}^*\|}, \quad y \geq 0. \tag{17}$$

The mean age of the system can be obtained from (**??**) as

$$\mathbb{E}[A] = -\mathbf{1}^T \, \mathsf{A}^{-1} \, \boldsymbol{\eta} = \frac{\|\mathsf{A}^{-1} \, \mathbf{x}^*\|}{\|\mathbf{x}^*\|}. \tag{18}$$

5.5 Age distribution of the compartments in steady state

Now suppose that compartment $j \in S$ has an unknown age density expressed by

$$x_j^* = x_j^* \int\limits_0^\infty f_{a_j}(y) \, dy.$$

In section **??** we have calculated the age of the regenerative process Z under the condition that its state equals j. This is the age that a traveling particle has when it is in compartment j. Consequently, using (**??**) and plugging in $\mathbf{x}^* = -\mathsf{A}^{-1}\,\mathbf{u}$, we get

$$f_{a_j}(y) = \frac{1}{x_j^*}\,\left(e^{y\,\mathsf{A}}\,\mathbf{u}\right)_j.$$

The mean age in compartment $j \in S$ is then, integrating by parts,

$$\int\limits_0^\infty y\, f_{a_j}(y)\, dy = \frac{1}{x_j^*}\,\left(\mathsf{A}^{-2}\,\mathbf{u}\right)_j = -\frac{1}{x_j^*}\,\left(\mathsf{A}^{-1}\,\mathbf{x}^*\right)_j.$$

Defining $\mathsf{X}^* := \operatorname{diag}(x_1^*,\dots,x_d^*)$, this gives a vector valued function of age densities in the compartments

$$\mathbf{f_a}(y) = (\mathsf{X}^*)^{-1}\,e^{y\,\mathsf{A}}\,\mathbf{u}, \quad y \geq 0,$$

and its expectation

$$\mathbb{E}[\mathbf{a}] = -(\mathsf{X}^*)^{-1}\,\mathsf{A}^{-1}\,\mathbf{x}^* \tag{19}$$

is the vector of mean ages in the compartments. We call it *mean age vector*.

If for $n \geq 1$ we denote by $\mathbf{a}^n$ the vector $(a_1^n,\dots,a_d^n)^T$. Multiple integration by parts leads to a formula for the vector whose components contain the nth moments of the compartment ages. It is given by

$$\mathbb{E}[\mathbf{a}^n] = (-1)^n\, n!\,(\mathsf{X}^*)^{-1}\,\mathsf{A}^{-n}\,\mathbf{x}^*.$$

Note that a computation of the average of the components of the mean age vector weighted by the respective steady state compartment content leads to the same equation for the mean system age as calculating the expected value of $\mathrm{PH}(\boldsymbol{\eta}, \mathsf{A})$ as it was done to obtain Eq. (**??**).

5.6 Transit time in steady state

For compartmental systems we can consider two types of transit times (**?**). The *forward transit time (FTT)* at time t_a is the time t a particle needs to travel through the system after it arrives at time t_a. The *backward transit time (BTT)* at time t_e gives the age y a particle has at the moment it leaves the system, that is, the time it needed to travel through the system given that it exits at time t_e. For an autonomous system in steady state we would expect the two types of transit time to coincide.

Obviously, the *FTT* is the absorption time of the absorbing continuous-time Markov chain X. As soon as the traveling particle leaves the system, X hits its absorbing state. Hence, the *FTT* is phase-type distributed with initial distribution $\boldsymbol{\beta}$ and transition rate matrix A. This makes its density

$$f_{FTT}(t) = \mathbf{z}^T\,e^{t\,\mathsf{A}}\,\boldsymbol{\beta}, \quad t \geq 0$$

and its mean

$$\mathbb{E}[FTT] = \|\mathsf{A}^{-1}\,\boldsymbol{\beta}\| = \frac{\|\mathbf{x}^*\|}{\|\mathbf{u}\|}. \tag{20}$$

For the BTT we can consider the above constructed renewal process. The BTT is the age of Z at the time of a restart. Hence, it follows the same distribution as the cycle length and that is identically distributed to the absorption time. So from the probabilistic point of view, FTT and BTT are obviously the same. But we can also calculate the density of the BTT as a weighted average of ages of particles that are leaving the system:

$$f_{BTT}(y) = \frac{1}{\|\mathbf{r}\|} \sum_{j \in S} r_j\, f_{a_j}(y) = \frac{1}{\|\mathbf{u}\|} \sum_{j \in S} \frac{z_j\,x_j^*}{x_j^*} \left(e^{y\,\mathsf{A}}\,\mathbf{u}\right)_j = \mathbf{z}^T\,e^{y\,\mathsf{A}}\,\boldsymbol{\beta}, \quad y \geq 0.$$

We can see that intuition does not betray us: FTT and BTT coincide in steady state. Furthermore, the transit time depends neither on the arrival time t_a nor on the exit time t_e.

6 Application to carbon cylce models

6.1 A linear autonomous compartmental system in steady state

As an example, we consider the global carbon-cycle model introduced by **?**, for which **?** numerically calculated age and transit time distributions using the impulse response function approach. This model is therefore a good test-case for our derivations. It comprises five compartments: non-woody tree parts x_1, woody tree parts x_2, ground vegetation x_3, detritus/decomposers x_4, and active soil carbon x_5. The model is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t) = \mathsf{A}\,\mathbf{x}(t) + \mathbf{u}, \quad t > 0, \tag{21}$$

where we use the input vector $\mathbf{u} = \begin{pmatrix} 77\ 0\ 36\ 0\ 0 \end{pmatrix}^T$ and the compartmental matrix

$$\mathsf{A} = \begin{pmatrix} -2.081 & 0 & 0 & 0 & 0 \\ 0.8378 & -0.0686 & 0 & 0 & 0 \\ 0 & 0 & -0.5217 & 0 & 0 \\ 0.5676 & 0.0322 & 0.1739 & -0.5926 & 0 \\ 0 & 0.004425 & 0.087 & 0.037 & -0.009813 \end{pmatrix}.$$

The numbers are taken exactly as in **?**. We also refer to that paper for comparison of the results.

The input vector is expressed in units of $\mathrm{Pg\,C\,yr}^{-1}$, and the fractional transfer coefficients in units of yr^{-1}. Because A is a lower triangular matrix with diagonal entries different from zero, it is nonsingular and we can apply all the earlier obtained formulas to system (**??**). The steady state vector of carbon content is

$$\mathbf{x}^* = -\mathsf{A}^{-1}\,\mathbf{u} = \begin{pmatrix} 37.0 & 451.89 & 69.01 & 80.24 & 1118.12 \end{pmatrix}^T \text{Pg C},$$

514 which results in a vector for the fractions of total carbon storage equal to

$$\begin{pmatrix} 2.11 & 25.73 & 3.93 & 4.57 & 63.66 \end{pmatrix}^T \text{\%}.$$

515 The respiration vector in steady state is

$$\begin{pmatrix} 25.0 & 14.45 & 18.0 & 44.58 & 10.97 \end{pmatrix}^T \text{Pg C yr}^{-1}.$$

516 That gives fractions of total respirations as

$$\begin{pmatrix} 22.12 & 12.79 & 15.93 & 39.45 & 9.71 \end{pmatrix}^T \text{\%}.$$

517 The transit time T is phase-type distributed with probability density function
518 (Fig. **??**)

$$f_T(t) = 0.31e^{-2.081t} - 0.3e^{-0.5926t} + 0.52e^{-0.5217t} + 0.018e^{-0.0686t} + 0.001e^{-0.009813t}.$$

519 The expected value $\mathbb{E}[T] = 15.54\,\text{yr}$, is almost identical to the value found by
520 **?** (15.58 yr); and the standard deviation is $\sigma_T = 44.95\,\text{yr}$.
521   Furthermore, we have the system age density given by

$$f_A(y) = 0.0096e^{-2.081y} - 0.033e^{-0.5926y} + 0.064e^{-0.5217y} + 0.017e^{-0.0686y} + 0.0066e^{-0.009813y}.$$

522 Its expectation $\mathbb{E}[A] = 72.78\,\text{yr}$ is also similar to that reported by **?** (72.82 yr),
523 with standard deviation $\sigma_A = 94.16\,\text{yr}$.
524   We can also calculate the vector that contains the probability density func-
525 tions for the age in the compartments as

$$\mathbf{f_a}(y) = \begin{pmatrix} 2.081 & 0 & 0 & 0 & 0 \\ -0.0709 & 0.0709 & 0 & 0 & 0 \\ 0 & 0 & 0.5217 & 0 & 0 \\ -0.3573 & 0.0245 & 1.1 & -0.7676 & 0 \\ 0.0005 & -0.0033 & -0.0112 & 0.0035 & 0.0104 \end{pmatrix} \begin{pmatrix} e^{-2.081y} \\ e^{-0.0686y} \\ e^{-0.5217y} \\ e^{-0.5926y} \\ e^{-0.009813y} \end{pmatrix},$$

526 from which we can get the mean age vector

$$\mathbb{E}[\mathbf{a}] = \begin{pmatrix} 0.48 & 15.06 & 1.92 & 6.99 & 107.59 \end{pmatrix}^T \text{yr}$$

527 and the standard deviation vector

$$\boldsymbol{\sigma_a} = \begin{pmatrix} 0.48 & 14.59 & 1.92 & 10.55 & 102.37 \end{pmatrix}^T \text{yr}.$$

528   From these density functions we can plot the system's and the compart-
529 ments' content with respect to their age (Fig. **??**), which give useful informa-
530 tion about the range of ages for each compartment and how they contribute
531 to the system-age distribution. In comparison to the results of **?**, our approach
532 not only provides mean values for ages and transit times, but also exact formu-
533 las for their probability distribution. In their approach, these authors obtained
534 results that depended on the simulation time and therefore included numerical
535 errors, something we can easily avoid.

536   6.2 A nonlinear autonomous compartmental system in steady state

537   Let us consider the nonlinear autonomous compartmental system

$$\frac{\mathrm{d}}{\mathrm{d}\,t}\mathbf{x}(t) = B(\mathbf{x}(t))\,\mathbf{x}(t) + \mathbf{u}, \quad t > 0, \tag{22}$$

538   where $B : \mathbb{R}^d \to \mathbb{R}^{d \times d}$ is a matrix-valued mapping. In this setup the fractional
539   transfer coefficients are not constant but depend on the system's content.
540   Consequently, the transition rates of the traveling particle are not constant.
541   Let us now assume that system (??) is in a steady state $\mathbf{x^*}$. From $\mathrm{d}\,\mathbf{x^*}/\mathrm{d}t =$
542   0 follows that the compartment content x_j do not change and the mapping B
543   turns into a matrix A with constant coefficients.
544   Hence, assuming the nonlinear autonomous compartmental system (??)
545   is in a steady state, we can treat it as a linear autonomous compartmental
546   system and apply the entire theory developed for those systems.
        As an example, consider the nonlinear two-compartment carbon-cycle model
        described by ?:

$$\frac{\mathrm{d}}{\mathrm{d}\,t} \begin{pmatrix} C_s \\ C_b \end{pmatrix} = \begin{pmatrix} -\lambda(\mathbf{x}) & \mu_b \\ \varepsilon\lambda(\mathbf{x}) & -\mu_b \end{pmatrix} \begin{pmatrix} C_s \\ C_b \end{pmatrix} + \begin{pmatrix} F_{npp} \\ 0 \end{pmatrix}.$$

547   Here, B depends on $\mathbf{x} = \begin{pmatrix} C_s & C_b \end{pmatrix}^T$ through λ's dependence on $\mathbf{x}$, which is
548   given by

$$\lambda(\mathbf{x}) = \frac{C_b V_s}{C_s + K_s}. \tag{23}$$

549   Steady state formulas for the compartment content can be computed as

$$C_s^* = \frac{K_s}{\frac{V_s \varepsilon}{\mu_b} - 1} \quad \text{and} \quad C_b^* = \frac{F_{npp}}{\mu_b \left(-1 + \frac{1}{\varepsilon}\right)}.$$

550   From ? we take the parameter values $F_{npp} = 345\,\mathrm{g\,C\,m^{-2}\,yr^{-1}}$, $\mu_b =$
551   $4.38\,\mathrm{yr^{-1}}$, $\varepsilon = 0.39$, and $K_s = 53954.83\,\mathrm{g\,C\,m^{-2}}$. Since the description of V_s
552   is missing in the original publication, we chose it to be equal to $59.13\,\mathrm{yr^{-1}}$ to
553   approximately meet the given steady state compartment content

$$C_s^* = 12650\,\mathrm{g\,C\,m^{-2}} \quad \text{and} \quad C_b^* = 50.36\,\mathrm{g\,C\,m^{-2}}.$$

554   With the given parameters, the steady state transfer matrix $\mathsf{A} = B(\mathbf{x^*})$ and
555   the input vector $\mathbf{u}$ are given by

$$\mathsf{A} = \begin{pmatrix} -0.0447 & 4.38 \\ 0.0174 & -4.38 \end{pmatrix} \quad \text{and} \quad \mathbf{u} = \begin{pmatrix} 345 \\ 0 \end{pmatrix}.$$

556   Our approach allows us to calculate mean transit time and mean ages together
557   with the according densities for different values of the model's parameters
558   (Fig. ??). In particular, we explore here the effect of different values of the
559   parameter ε on the steady state ages and transit times. This parameter controls
560   the proportion of carbon that is transferred from the substrate C_s to the

microbial biomass compartment C_b, and it is commonly referred to as the *carbon use efficiency*. Interestingly, if the carbon use efficiency ε increases, the mean transit time and mean ages of the model at steady state decrease (Fig. **??**), a behavior that at first glance appears counterintuitive. It can be explained by two opposing effects. On the one hand, an increase of carbon use efficiency keeps a higher fraction of carbon in the system due to lower respiration. This has an increasing effect on the transit time. On the other hand, a higher carbon use efficiency ε implies a lower steady state content of compartment C_s and a higher one of compartment C_b. Consequently, from (**??**) we get an increasing value of

$$\lambda(\mathbf{x}^*) = \frac{F_{npp} V_s}{\mu_b \left(-1 + \frac{1}{\varepsilon}\right) \left(K_s + \frac{K_s}{\frac{V_s \varepsilon}{\mu_b} - 1}\right)}.$$

This value is the process rate of the C_s compartment in steady state. The higher it is, the faster the particles travel through the system. The latter effect wins here and we see a decrease in the transit time.

The graph of the mean system age for this model with $\varepsilon = 0.39$ (Fig. **??**) lies directly on the one of the mean transit time. The huge difference in the compartments' steady state content causes very little difference in the initial probability vectors $\boldsymbol{\beta}$ and $\boldsymbol{\eta}$ for the traveling particle. This results in very similar distributions of transit time and system age.

7 Discussion

We have derived formulas for the densities, higher order moments, and first moments for the transit time, the system age, and the age vector. They can be found in different places in this manuscript. For convenience, Table **??** provides a quick overview. We further provide a small Python package for linear autonomous compartmental models, called LAPM, on `https://github.com/goujou/LAPM`. It deals with symbolic and numerical computations of the formulas given in Table **??**.

7.1 Relation to earlier results

Over many years compartmental models with fixed coefficients have been studied (**???**) by mainly investigating the mean age of particles in the system and the mean transit time of particles. The topic became of increasing interest again with the need for modeling the terrestrial carbon cycle.

The general situation of linear nonautonomous compartmental models was investigated by **?**. For the special case of autonomous systems they provide the formulas (**??**), (**??**), (**??**) for the mean age vector, mean system age, and mean transit time, respectively. Our probabilistic approach shows that these

formulas are just expected values of corresponding random variables with according probability distributions $\mathbf{f_a}(y) = -(\mathsf{X}^*)^{-1} e^{y\,\mathsf{A}}\,\mathbf{u}$, $A \sim \mathrm{PH}(\boldsymbol{\eta}, \mathsf{A})$, and $T \sim \mathrm{PH}(\boldsymbol{\beta}, \mathsf{A})$. In particular, the terms *mean system age* and *mean transit time* refer to the expected absorption times of related continuous-time Markov chains.

A general formula for the moments of the distribution of the transit time in a compartmental system without external inputs was already given by **?**. He also states a formula for the backward transit time density in this situation.

? (also **??**) calculated a so-called *transfer function* or *input response function* Ψ that describes the relative contribution of the input at time $t - T$ to the output at time t. They furthermore give a *survivor function*

$$\tilde{A}(t) = \int\limits_{t}^{\infty} \Psi(T)\, dT,$$

that describes the probability of the particle to remain in the system for at least a time t. In our setup this means

$$\tilde{A}(t) = 1 - F_T(t),$$

which leads us by

$$1 - F_T(t) = \int\limits_{t}^{\infty} f_T(t)\, \mathrm{d}t$$

to the conclusion that the input response function Ψ is nothing else than the probability density function f_T of a phase-type distribution $\mathrm{PH}(\boldsymbol{\beta}, \mathsf{A})$ given by (**??**). They also computed an *age density distribution*

$$\varphi(t) = \frac{\tilde{A}(t)}{\int\limits_{0}^{\infty} \tilde{A}(\tau)\, \mathrm{d}\tau} = \frac{\tilde{A}(t)}{\overline{T}},$$

which is exactly the probability density function f_A of the system age given by (**??**). This becomes obvious from (**??**) which perfectly relates the distributions of transit time and system age.

For special cases of simple compartmental systems, they further provide formulas for the densities and means of transit time and system age. These formulas were obtained by simulating an impulsive input and considering the impulse response in the Laplace space. It turns out that those formulas are special cases of the general shape of probability density functions of phase-type. See Appendix **??** for examples.

7.2 Singular compartmental matrices

In this paper, we considered only nonsingular compartmental or transition rate matrices, because they ensure that every particle that enters the system will eventually leave it. It can be shown (**?**) that the transit time of a continuous-time Markov chain with singular transition rate matrix does not lead to a certainly finite transit time. In this case the system contains traps from which particles are not able to get out once they have entered them. So they will remain in the system forever. Although this seems unrealistic, there are successful carbon-cycle models with this structure.

A famous example is the RothC model (**?**). It contains an inert compartment that has neither inputs nor outputs, but it has a positive constant content. To apply our stochastic framework to this model we need to cut this compartment out of the system and look at the steady state of the remaining, smaller system. The transit time of newly arriving particles will not be affected. The age distribution of the system, however, will depend on time since the particles in the inert compartment become older and older. A detailed analysis of how to treat systems with traps can be found in **?** and references therein.

7.3 Stochastic versus deterministic approach

As noted by **?**, in modeling ecological systems, there are powerful arguments for the use of stochastic processes. Even if nature is completely deterministic, ecological system are too complex for our complete theoretical understanding or descriptive tools. We can allow for this lack of complete knowledge by using probabilistic methods. One of the first authors to integrate stochastic behavior in biological models was **?**. **?** gives an early review on stochastic compartmental models with ecological applications. He shows how compartmental models can be considered as a deterministic representation of an underlying stochastic process by establishing the link to continuous-time Markov chains. While **?** elaborates on this link and applies it to a liver model, we connect the continuous-time Markov chain approach to the theory of phase-type distributions (**?**) and focus on transit time and ages.

The deterministic theory describes the average or ideal behavior of a system, whereas stochastic models can examine the deviations from such ideal or average behavior. The stochastic approach gives us the opportunity to think of key quantities as random variables. Knowledge of their probability distributions gives us in hand a whole new set of tools not only to study the mean, but also higher order moments of such random variables. A high variance of the transit time, for example, tells us that it is not the usual behavior of particles to travel though the system as long as indicated by the mean transit time. There is rather a huge variety of different transit times depending on the particular path taken by the particle. Long tails of the distribution mean that there are in fact particles that need very long to travel through the system and

hence there are very old masses in the system. With the probability density functions in hand we can further think of quantiles and confidence intervals of transit time and system age, something not possible with the deterministic approach.

Linear autonomous compartmental carbon-cycle models relate to continuous-time Markov chains with time invariant transition probabilities. If we turn to nonautonomous models (**?**), we would have to consider time variant Markov chains in the probabilistic approach. In this situation the Kolmogorov equations govern the evolution of the transition probabilities. There is plenty of theory available on continuous-time Markov chains (**??**), and much of it might be applied to carbon-cycle models.

8 Conclusions

We modeled open linear autonomous compartmental systems by absorbing continuous-time Markov chains and showed that the behavior of those systems is governed by the phase-type distribution. It applies for the transit time as well as for the system age, only with different initial distributions. This knowledge provides us with simple general formulas for the densities, means, and higher order moments of transit time and system age. Furthermore, it explains the underlying structure of the intrinsic connection between system age and transit time (**?**).

Our approach also revealed the potential and opportunities of expressing deterministic compartmental models as stochastic processes. We showed how important system diagnostics in reservoir theory such as system ages and transit times have analogous counterparts in probabilistic terms. The theory of stochastic processes can further help to address important questions in reservoir theory with applications in other fields such as the global biogeochemical cycles.

9 Acknowledgements

A Proof of Theorem ??

We cannot draw a random time uniformly from the positive half-line, so we draw τ uniformly from the interval $[0, L]$ for some $L > 0$ and illustrate τ's dependence on L by writing τ_L. We aim at finding $\mathbb{P}(Y_{\tau_L} = 1)$ and letting L tend to infinity. The existence of a random variable Y_τ according to the limiting distribution is then guaranteed by Skorokhods representation theorem (**?**).

Note that the random variable Y_τ is stochastic in two ways: The value of Y at a deterministic time is stochastic and furthermore the time τ itself is stochastic. To overcome

this difficulty we condition the probability $\mathbb{P}(Y_{\tau_L} = 1)$ to $\tau_L = t$ and use the law of total probability to rewrite it as

$$\mathbb{P}(Y_{\tau_L} = 1) = \int\limits_0^\infty \mathbb{P}(Y_{\tau_L} = 1 \mid \tau_L = t)\, f_{\tau_L}(t)\, \mathrm{d}t,$$

where $f_{\tau_L}(t) = \mathbb{1}_{[0,L]}(t)\, 1/L$ denotes the probability density function of the uniform distribution on $[0, L]$. We plug it in to get

$$\mathbb{P}(Y_{\tau_L} = 1) = \frac{1}{L} \int\limits_0^L \mathbb{P}(Y_t = 1)\, \mathrm{d}t, \quad \text{for } y \geq 0.$$

Let $p(t) := \mathbb{P}(Y_t = 1)$. If we knew the existence of $\lim_{t \to \infty} p(t)$, we could conclude

$$\lim_{L \to \infty} \mathbb{P}(Y_{\tau_L} = 1) = \lim_{L \to \infty} \frac{1}{L} \int\limits_0^L p(t)\, \mathrm{d}t = \lim_{t \to \infty} p(t). \qquad (24)$$

We are left with the problem of calculating $\lim_{t \to \infty} p(t)$. To solve it we need a major result from renewal theory. Recall that f_T denotes the common probability density function of the cycle durations.

Theorem 2 (Key renewal theorem) *Suppose that the function z in the so-called* general renewal equation

$$p(t) = z(t) + \int\limits_0^t p(t - x)\, f_T(x)\, \mathrm{d}x, \quad t \geq 0, \qquad (25)$$

is directly Riemann integrable. Then

$$\lim_{t \to \infty} p(t) = \frac{1}{\mathbb{E}[T]} \int\limits_0^\infty z(\tau)\, \mathrm{d}\tau.$$

The proof and the rather technical definition of direct Riemann integrability can be found in **?**, Sect. V.4.

To apply the key renewal theorem to p, we have to show that p satisfies the renewal equation (**??**) for some directly Riemann integrable function z. To this end we use a standard technique from renewal theory and condition on the event that the first restart occurs at time $T_1 = x$. Recall that T_1 is distributed according to the probability density function f_T. We see

$$p(t) = \int\limits_0^\infty \mathbb{P}(Y_t = 1 \mid T_1 = x)\, f_T(x)\, \mathrm{d}x.$$

Now we split this integral into a period before and a period after t to get

$$p(t) = \int\limits_0^t \mathbb{P}(Y_t = 1 \mid T_1 = x)\, f_T(x)\, \mathrm{d}x + \int\limits_t^\infty \mathbb{P}(Y_t = 1 \mid T_1 = x)\, f_T(x)\, \mathrm{d}x.$$

The second integral describes the probability of Y being on during the first cycle, since $t \leq T_1$, and is equal to $z(t) := \mathbb{P}(Y_t^1 = 1)$. The conditional probability $\mathbb{P}(Y_t = 1 \mid T_1 = x)$ under the first integral equals $\mathbb{P}(Y_{t-x} = 1) = p(t - x)$, since the process restarts at the end $T_1 = x$ of the first cycle. Hence,

$$p(t) = \int\limits_0^t p(t - x)\, f_T(x)\, \mathrm{d}x + z(t), \quad t \geq 0,$$

satisfies the general renewal equation (??). Furthermore, the function z can be shown to be directly Riemann integrable. Therefore, the key renewal theorem states that

$$\lim_{t \to \infty} p(t) = \frac{1}{E[T]} \int\limits_0^\infty \mathbb{P}(Y_t^1 = 1) \, \mathrm{d}t. \tag{26}$$

The integral can be rewritten to

$$\int\limits_0^\infty \mathbb{P}(Y_t^1 = 1) \, \mathrm{d}t = \int\limits_0^\infty \mathbb{E}\mathbb{1}_{\{Y_t^1 = 1\}}(t) \, \mathrm{d}t = \mathbb{E} \int\limits_0^\infty \mathbb{1}_{\{Y_t^1 = 1\}}(t) \, \mathrm{d}t$$

and hence describes the expected time of Y being on during the first cycle. We consider the equality of the limits in Eq. (??) and have finished the poof.

B Examples

B.1 One single compartment

We consider the one-compartment model represented by the ordinary linear differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} x(t) = -\lambda \, x(t) + u, \quad t > 0,$$

for $\lambda > 0$. In this simplest possible framework we have $\mathsf{A} = -\lambda$, $\mathbf{z} = \lambda$, $\mathsf{A}^{-1} = -1/\lambda$, and $\boldsymbol{\beta} = 1$. The according phase-type distribution is just the exponential distribution. The cumulative distribution function of the transit time T is

$$F_T(t) = 1 - e^{-\lambda \, t}, \quad t \geq 0,$$

its probability density function is

$$f_T(t) = \lambda \, e^{-\lambda \, t}, \quad t \geq 0,$$

and the expected absorption time or mean transit time is $\mathbb{E}[T] = 1/\lambda$. The mean age vector $\mathbf{a}$ coincides with the system age A and its probability density function is

$$f_A(y) = \lambda \, e^{-\lambda \, y}, \quad y \geq 0,$$

which leads to the mean age of $\mathbb{E}[A] = 1/\lambda$. The fact that transit time and age have the same distribution reflects the memorylessness of the exponential distribution.

B.2 Two compartments without feedback

A more interesting example is a two-compartment system given by

$$\frac{\mathrm{d}}{\mathrm{d}t} x_1(t) = -\lambda_1 \, x_1(t) + u_1,$$

$$\frac{\mathrm{d}}{\mathrm{d}t} x_2(t) = \alpha \, \lambda_1 \, x_1(t) - \lambda_2 \, x_2(t) + u_2,$$

with $\lambda_1 > 0$, $\lambda_2 > 0$, and $0 \leq \alpha \leq 1$. We are given

$$\mathsf{A} = \begin{pmatrix} -\lambda_1 & 0 \\ \alpha\lambda_1 & -\lambda_2 \end{pmatrix}, \quad \mathsf{A}^{-1} = \begin{pmatrix} -\frac{1}{\lambda_1} & 0 \\ -\frac{\alpha}{\lambda_2} & -\frac{1}{\lambda_2} \end{pmatrix},$$

744   and $\mathbf{z} = \left(-\alpha\lambda_1 + \lambda_1 \ \lambda_2\right)^T, \quad \mathbf{u} = \left(u_1 \ u_2\right)^T, \ \boldsymbol{\beta} = \left(\frac{u_1}{u_1+u_2} \ \frac{u_2}{u_1+u_2}\right)^T.$

745   If $\lambda_1 \neq \lambda_2$ the matrix exponential is

$$e^{t\,\mathsf{A}} = \begin{pmatrix} e^{-\lambda_1 t} & 0 \\ \frac{\alpha\lambda_1}{\lambda_1-\lambda_2}\left(e^{-\lambda_2 t} - e^{-\lambda_1 t}\right) & e^{-\lambda_2 t} \end{pmatrix},$$

746   and the cumulative distribution function of the transit time T is

$$F_T(t) = -\frac{\alpha\lambda_1 u_1 \left(e^{-\lambda_2 t} - e^{-\lambda_1 t}\right)}{(\lambda_1 - \lambda_2)(u_1 + u_2)} - \frac{u_1 e^{-\lambda_1 t}}{u_1 + u_2} - \frac{u_2 e^{-\lambda_2 t}}{u_1 + u_2} + 1.$$

747   The probability density function is

$$f_T(t) = \frac{\lambda_2 u_2 e^{-\lambda_2 t}}{u_1 + u_2} + \frac{u_1}{u_1 + u_2}\left(\frac{\alpha\lambda_1\lambda_2}{\lambda_1 - \lambda_2}\left(e^{-\lambda_2 t} - e^{-\lambda_1 t}\right) + (-\alpha\lambda_1 + \lambda_1)e^{-\lambda_1 t}\right)$$

748   and the expected absorption time

$$\mathbb{E}[T] = \frac{u_1\left(\frac{\alpha}{\lambda_2} + \frac{1}{\lambda_1}\right)}{u_1 + u_2} + \frac{u_2}{\lambda_2(u_1 + u_2)}.$$

749   For the age distribution we need to calculate the steady state solution and its normalized
750   version first:

$$\mathbf{x}^* = \mathsf{A}^{-1}\,\mathbf{u} = \left(\frac{u_1}{\lambda_1} \ \frac{\alpha u_1}{\lambda_2} + \frac{u_2}{\lambda_2}\right)^T, \quad \boldsymbol{\eta} = \frac{\mathbf{x}^*}{\|\mathbf{x}^*\|} = \left(\frac{u_1}{\lambda_1\left(\frac{\alpha u_1}{\lambda_2} + \frac{u_2}{\lambda_2} + \frac{u_1}{\lambda_1}\right)} \ \frac{\frac{\alpha u_1}{\lambda_2} + \frac{u_2}{\lambda_2}}{\frac{\alpha u_1}{\lambda_2} + \frac{u_2}{\lambda_2} + \frac{u_1}{\lambda_1}}\right)^T.$$

Now we use that $A \sim \mathrm{PH}(\boldsymbol{\eta}, \mathsf{A})$ and obtain, for $y \geq 0$,

$$F_A(y) = -\frac{\alpha u_1\left(e^{-\lambda_2 y} - e^{-\lambda_1 y}\right)}{(\lambda_1 - \lambda_2)\left(\frac{\alpha u_1}{\lambda_2} + \frac{u_2}{\lambda_2} + \frac{u_1}{\lambda_1}\right)} - \frac{\left(\frac{\alpha u_1}{\lambda_2} + \frac{u_2}{\lambda_2}\right)e^{-\lambda_2 y}}{\frac{\alpha u_1}{\lambda_2} + \frac{u_2}{\lambda_2} + \frac{u_1}{\lambda_1}} + 1 - \frac{u_1 e^{-\lambda_1 y}}{\lambda_1\left(\frac{\alpha u_1}{\lambda_2} + \frac{u_2}{\lambda_2} + \frac{u_1}{\lambda_1}\right)},$$

$$f_A(y) = \frac{\lambda_2\left(\frac{\alpha u_1}{\lambda_2} + \frac{u_2}{\lambda_2}\right)e^{-\lambda_2 y}}{\frac{\alpha u_1}{\lambda_2} + \frac{u_2}{\lambda_2} + \frac{u_1}{\lambda_1}} + \frac{u_1}{\lambda_1\left(\frac{\alpha u_1}{\lambda_2} + \frac{u_2}{\lambda_2} + \frac{u_1}{\lambda_1}\right)}\left(\frac{\alpha\lambda_1\lambda_2}{\lambda_1 - \lambda_2}\left(e^{-\lambda_2 y} - e^{-\lambda_1 y}\right) + (-\alpha\lambda_1 + \lambda_1)e^{-\lambda_1 y}\right),$$

$$\mathbb{E}[A] = \frac{\frac{\alpha u_1}{\lambda_2} + \frac{u_2}{\lambda_2}}{\lambda_2\left(\frac{\alpha u_1}{\lambda_2} + \frac{u_2}{\lambda_2} + \frac{u_1}{\lambda_1}\right)} + \frac{u_1\left(\frac{\alpha}{\lambda_2} + \frac{1}{\lambda_1}\right)}{\lambda_1\left(\frac{\alpha u_1}{\lambda_2} + \frac{u_2}{\lambda_2} + \frac{u_1}{\lambda_1}\right)}.$$

751   The vector valued function containing the age densities of the compartments is

$$\mathbf{f_a}(y) = \left(\lambda_1 e^{-\lambda_1 y} \ \frac{\alpha\lambda_1\lambda_2 u_1\left(e^{-\lambda_2 y} - e^{-\lambda_1 y}\right)}{(\lambda_1 - \lambda_2)(\alpha u_1 + u_2)} + \frac{\lambda_2 u_2 e^{-\lambda_2 y}}{\alpha u_1 + u_2}\right)^T.$$

752   This leads to the following mean age vector:

$$\mathbb{E}[\mathbf{a}] = \left(\frac{1}{\lambda_1} \ \frac{\alpha u_1}{\lambda_1(\alpha u_1 + u_2)} + \frac{\frac{\alpha u_1}{\lambda_2} + \frac{u_2}{\lambda_2}}{\alpha u_1 + u_2}\right)^T.$$

753   *B.2.1 Serial compartments - Hypoexponential distribution*

If $u_2 = 0$ and $\alpha = 1$, the particle enters the system in compartment 1 and must travel
through compartment 2 before absorption. This leads to the transit time T being hypoex-
ponentially distributed. That is, T is distributed like the sum of two independent exponential

distributions:

$$F_T(t) = -\frac{\lambda_1}{\lambda_1 - \lambda_2}\left(e^{-\lambda_2 t} - e^{-\lambda_1 t}\right) + 1 - e^{-\lambda_1 t},$$

$$f_T(t) = \frac{\lambda_1 \lambda_2}{\lambda_1 - \lambda_2}\left(e^{-\lambda_2 t} - e^{-\lambda_1 t}\right),$$

$$\mathbb{E}[T] = \frac{1}{\lambda_2} + \frac{1}{\lambda_1}.$$

The steady state solution and its normalized version are

$$\mathbf{x^*} = \left(\frac{u_1}{\lambda_1} \; \frac{u_1}{\lambda_2}\right)^T \quad \text{and} \quad \boldsymbol{\eta} = \left(\frac{u_1}{\lambda_1\left(\frac{u_1}{\lambda_2} + \frac{u_1}{\lambda_1}\right)} \; \frac{u_1}{\lambda_2\left(\frac{u_1}{\lambda_2} + \frac{u_1}{\lambda_1}\right)}\right)^T.$$

For the system age we get

$$F_A(y) = -\frac{u_1\left(e^{-\lambda_2 y} - e^{-\lambda_1 y}\right)}{(\lambda_1 - \lambda_2)\left(\frac{u_1}{\lambda_2} + \frac{u_1}{\lambda_1}\right)} + 1 - \frac{u_1 e^{-\lambda_2 y}}{\lambda_2\left(\frac{u_1}{\lambda_2} + \frac{u_1}{\lambda_1}\right)} - \frac{u_1 e^{-\lambda_1 y}}{\lambda_1\left(\frac{u_1}{\lambda_2} + \frac{u_1}{\lambda_1}\right)},$$

$$f_A(y) = \frac{\lambda_2 u_1\left(e^{-\lambda_2 y} - e^{-\lambda_1 y}\right)}{(\lambda_1 - \lambda_2)\left(\frac{u_1}{\lambda_2} + \frac{u_1}{\lambda_1}\right)} + \frac{u_1 e^{-\lambda_2 y}}{\frac{u_1}{\lambda_2} + \frac{u_1}{\lambda_1}},$$

$$\mathbb{E}[A] = \frac{u_1}{\lambda_2^2\left(\frac{u_1}{\lambda_2} + \frac{u_1}{\lambda_1}\right)} + \frac{u_1\left(\frac{1}{\lambda_2} + \frac{1}{\lambda_1}\right)}{\lambda_1\left(\frac{u_1}{\lambda_2} + \frac{u_1}{\lambda_1}\right)}.$$

For $u_1 = 1$ this turns into

$$F_A(y) = -\frac{\alpha\left(e^{-\lambda_2 y} - e^{-\lambda_1 y}\right)}{(\lambda_1 - \lambda_2)\left(\frac{\alpha}{\lambda_2} + \frac{1}{\lambda_1}\right)} - \frac{\alpha e^{-\lambda_2 y}}{\lambda_2\left(\frac{\alpha}{\lambda_2} + \frac{1}{\lambda_1}\right)} + 1 - \frac{e^{-\lambda_1 y}}{\lambda_1\left(\frac{\alpha}{\lambda_2} + \frac{1}{\lambda_1}\right)},$$

$$f_A(y) = \frac{\alpha e^{-\lambda_2 y}}{\frac{\alpha}{\lambda_2} + \frac{1}{\lambda_1}} + \frac{1}{\lambda_1\left(\frac{\alpha}{\lambda_2} + \frac{1}{\lambda_1}\right)}\left(\frac{\alpha \lambda_1 \lambda_2}{\lambda_1 - \lambda_2}\left(e^{-\lambda_2 y} - e^{-\lambda_1 y}\right) + (-\alpha\lambda_1 + \lambda_1)\, e^{-\lambda_1 y}\right),$$

$$\mathbb{E}[A] = \frac{\alpha}{\lambda_2^2\left(\frac{\alpha}{\lambda_2} + \frac{1}{\lambda_1}\right)} + \frac{1}{\lambda_1}.$$

The vector valued function containing the age densities of the compartments is

$$\mathbf{f_a}(y) = \left(\lambda_1 e^{-\lambda_1 y} \; \frac{\lambda_1 \lambda_2}{\lambda_1 - \lambda_2}\left(e^{-\lambda_2 y} - e^{-\lambda_1 y}\right)\right)^T,$$

which leads to the mean age vector $\mathbb{E}[\mathbf{a}] = \left(\frac{1}{\lambda_1} \; \frac{1}{\lambda_2} + \frac{1}{\lambda_1}\right)^T$.

If furthermore $\lambda_1 = \lambda_2$, then the hypoexponential distribution turns into an Erlang distribution which is the sum of two independent and identically distributed exponential distributions.

B.2.2 Parallel compartments - Hyperexponential distribution

In case of $\alpha = 0$ we get a purely parallel system and the transit time T is hyperexponentially distributed, that is, for $t \geq 0$,

$$F_T(t) = -\frac{u_1 e^{-\lambda_1 t}}{u_1 + u_2} - \frac{u_2 e^{-\lambda_2 t}}{u_1 + u_2} + 1,$$

$$f_T(t) = \frac{\lambda_1 u_1 e^{-\lambda_1 t}}{u_1 + u_2} + \frac{\lambda_2 u_2 e^{-\lambda_2 t}}{u_1 + u_2},$$

$$\mathbb{E}[T] = \frac{u_2}{\lambda_2(u_1 + u_2)} + \frac{u_1}{\lambda_1(u_1 + u_2)}.$$

761   The steady state solution and its normalized version are

$$\mathbf{x}^* = \left(\frac{u_1}{\lambda_1} \; \frac{u_2}{\lambda_2} \right)^T \quad \text{and} \quad \boldsymbol{\eta} = \left(\frac{u_1}{\lambda_1 \left(\frac{u_2}{\lambda_2} + \frac{u_1}{\lambda_1} \right)} \; \frac{u_2}{\lambda_2 \left(\frac{u_2}{\lambda_2} + \frac{u_1}{\lambda_1} \right)} \right)^T.$$

For the system age we get, for $t \geq 0$,

$$F_A(y) = 1 - \frac{u_2 e^{-\lambda_2 y}}{\lambda_2 \left(\frac{u_2}{\lambda_2} + \frac{u_1}{\lambda_1} \right)} - \frac{u_1 e^{-\lambda_1 y}}{\lambda_1 \left(\frac{u_2}{\lambda_2} + \frac{u_1}{\lambda_1} \right)},$$

$$f_A(y) = \frac{u_1 e^{-\lambda_1 y}}{\frac{u_2}{\lambda_2} + \frac{u_1}{\lambda_1}} + \frac{u_2 e^{-\lambda_2 y}}{\frac{u_2}{\lambda_2} + \frac{u_1}{\lambda_1}},$$

$$\mathbb{E}[A] = \frac{u_2}{\lambda_2^2 \left(\frac{u_2}{\lambda_2} + \frac{u_1}{\lambda_1} \right)} + \frac{u_1}{\lambda_1^2 \left(\frac{u_2}{\lambda_2} + \frac{u_1}{\lambda_1} \right)}.$$

762   The vector valued function containing the age densities of the compartments is $\mathbf{f_a}(y) =$
763     $\left(\lambda_1 e^{-\lambda_1 y} \; \lambda_2 e^{-\lambda_2 y} \right)^T$. This leads to the mean age vector $\mathbb{E}[\mathbf{a}] = \left(\frac{1}{\lambda_1} \; \frac{1}{\lambda_2} \right)^T$.

764   ## B.3 Two compartments with feedback

? considered also the simple two-compartment system with feedback

$$\frac{\mathrm{d}}{\mathrm{d}t} x_1(t) = -\lambda_1 \, x_1(t) + \lambda_2 \, x_2(t) + u_1,$$

$$\frac{\mathrm{d}}{\mathrm{d}t} x_2(t) = \alpha \, x_1(t) - \lambda_2 \, x_2(t),$$

765   in which mass enters and leaves the system only through the first compartment, but can in
766   between spend some time in the second compartment. The compartmental matrix and the
767   input vector are

$$\mathsf{A} = \begin{pmatrix} -\lambda_1 & \lambda_2 \\ \alpha \lambda_1 & -\lambda_2 \end{pmatrix} \quad \text{and} \quad \mathbf{u} = \begin{pmatrix} u_1 \\ 0 \end{pmatrix};$$

768   and ? provide the Laplacian of the density of the transit time and the system age. The
769   Laplacian of a PH$(\boldsymbol{\beta}, \mathsf{A})$ distribution is given by

$$L(\sigma) = \mathbf{z}^T \, (\sigma \, \mathsf{I} - \mathsf{A})^{-1} \, \boldsymbol{\beta}.$$

770   Consequently the Laplacian for the transit time is

$$L_T(\sigma) = \frac{\lambda_1 \, (\alpha - 1) \, (\lambda_2 + s)}{\alpha \lambda_1 \lambda_2 - (\lambda_1 + s) \, (\lambda_2 + s)},$$

771   and the Laplacian for the age is

$$L_A(\sigma) = \frac{\lambda_1 \lambda_2 \, (\alpha - 1) \, (\alpha \lambda_1 + \lambda_2 + s)}{(\alpha \lambda_1 + \lambda_2) \, (\alpha \lambda_1 \lambda_2 - (\lambda_1 + s) \, (\lambda_2 + s))}.$$

772   The expected values are given by

$$E[T] = -\frac{\alpha \lambda_1 + \lambda_2}{\lambda_1 \lambda_2 \, (\alpha - 1)},$$

773   and

$$E[A] = -\frac{\alpha \lambda_1 \, (\lambda_1 + \lambda_2) + \lambda_2 \, (\alpha \lambda_1 + \alpha \lambda_2 - \lambda_2 \, (\alpha - 1))}{\lambda_1 \lambda_2 \, (\alpha - 1) \, (\alpha \lambda_1 + \lambda_2)}.$$

C Tables and Figures

Table 1 Overview of derived formulas for open linear autonomous compartmental systems $\frac{\mathrm{d}}{\mathrm{d}\,t}\,\mathbf{x}(t) = \mathsf{A}\,\mathbf{x}(t) + \mathbf{u}$.

Metric	Density	nth moment	First moment
Transit time	$\mathbf{z}^T\,e^{t\,\mathsf{A}}\,\frac{\mathbf{u}}{\|\mathbf{u}\|}$	$(-1)^n\,n!\,\mathbf{1}^T\,\mathsf{A}^{-n}\,\frac{\mathbf{u}}{\|\mathbf{u}\|}$	$\dfrac{-\mathbf{1}^T\,\mathsf{A}^{-1}\,\frac{\mathbf{u}}{\|\mathbf{u}\|}}{\frac{\|\mathbf{x}^*\|}{\|\mathbf{u}\|}}$
System age	$\mathbf{z}^T\,e^{y\,\mathsf{A}}\,\frac{\mathbf{x}^*}{\|\mathbf{x}^*\|}$	$(-1)^n\,n!\,\mathbf{1}^T\,\mathsf{A}^{-n}\,\frac{\mathbf{x}^*}{\|\mathbf{x}^*\|}$	$\dfrac{-\mathbf{1}^T\,\mathsf{A}^{-1}\,\frac{\mathbf{x}^*}{\|\mathbf{x}^*\|}}{\frac{\|\mathsf{A}^{-1}\,\mathbf{x}^*\|}{\|\mathbf{x}^*\|}}$
Age vector	$(\mathsf{X}^*)^{-1}\,e^{y\,\mathsf{A}}\,\mathbf{u}$	$(-1)^n\,n!\,(\mathsf{X}^*)^{-1}\,\mathsf{A}^{-n}\,\mathbf{x}^*$	$-(\mathsf{X}^*)^{-1}\,\mathsf{A}^{-1}\,\mathbf{x}^*$

$\mathbf{z}^T = -\mathbf{1}^T\,\mathsf{A}$ is the row vector of release rates.
$\mathbf{x}^* = -\mathsf{A}^{-1}\,\mathbf{u}$ is the steady state vector.
$\mathsf{X}^* = \mathrm{diag}(x_1^*, x_2^*, \ldots, x_d^*)$ is the diagonal matrix comprising the components of the steady state vector.