International Institute for
Applied Systems Analysis
Schlossplatz 1
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342
Fax: +43 2236 71313
E-mail: publications@iiasa.ac.at
Web: www.iiasa.ac.at

Interim Report **IR-02-048**

On the Formulation and Analysis of General Deterministic Structured Population Models II. Nonlinear Theory

Odo Diekmann (o.diekmann@math.ruu.nl)
Mats Gyllenberg (matsgyl@utu.fi)
Haiyang Huang (hhywsg@bnu.ihep.ac.cn)
Markus Kirkilionis (markus.kirkilionis@iwr.uni-heidelberg.de)
J.A.J. Metz (metz@rulsfb.leidenuniv.nl)
Horst R. Thieme (h.thieme@asu.edu)

Approved by

Ulf Dieckmann (dieckman@iiasa.ac.at)
Project Leader, Adaptive Dynamics Network

June 2002

Contents

Abstract

This paper is as much about a certain modelling methodology, as it is about the constructive definition of future population states from a description of individual behaviour and an initial population state. The key idea is to build a nonlinear model in two steps, by explicitly introducing the environmental condition via the requirement that individuals are independent from one another (and hence equations are linear) when this condition is given (prescribed) as a function of time.

A linear physiologically structured population model is defined by two rules, one for reproduction and one for development and survival, both depending on the initial individual state and the prevailing environmental condition. In Part I we showed how one can constructively define future population state operators from these two ingredients.

A nonlinear model is a linear model together with a feedback law that describes how the environmental condition at any particular time depends on the population size and composition at that time. When applied to the solution of the linear problem, the feedback law yields a fixed point problem. This we solve constructively by means of the contraction mapping principle, for any given initial population state. Using subsequently this fixed point as input in the linear population model, we obtain a population semiflow. We then say that we solved the nonlinear problem.

The paper is organized in a top-down spirit: We describe a general abstract setting first and then specialise, while becoming more technical.

The results are not restricted to a single population but also cover the interaction (including predation) of several structured (and unstructured) populations.

About the Authors

Odo Diekmann
CWI, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
and
Institute of Evolutionary and Ecological Sciences
University of Leiden
Kaiserstraat 63, 2311 GP Leiden, The Netherlands

Mats Gyllenberg
Department of Mathematics, University of Turku
FIN-20500 Turku, Finland

Haiyang Huang
Department of Mathematics, Beijing Normal University
Beijing 100875, P.R. of China

Markus Kirkilionis
Interdisziplinäres Inst. f. wiss. Rechnen
University of Heidelberg
Im Neuenheimer Feld 368, 69120 Heidelberg, Germany

J.A.J. Metz
Institute of Evolutionary and Ecological Sciences
University of Leiden
Kaiserstraat 63, 2311 GP Leiden, The Netherlands
and
CWI, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
and
Adaptive Dynamics Network
International Institute for Applied Systems Analysis
A-2361 Laxenburg, Austria

Horst R. Thieme
Department of Mathematics, Arizona State University
Tempe, AZ 85287-1804, USA

Acknowledgement

The Lorentz Center of Leiden University housed five of the six authors during the week July 5-9, 1999 and the Academy of Finland financed a meeting of three of the authors during November 11-15, 1999. As these two meetings were instrumental in getting the paper finished, we are most obliged. The research of H.H. has been supported by the National Natural Science Foundation of China and the research of H.R.T. by NSF grants DMS-9403884 and DMS-9706787.

On the Formulation and Analysis of General Deterministic Structured Population Models
II. Nonlinear Theory

Odo Diekmann
Mats Gyllenberg
Haiyang Huang
Markus Kirkilionis
J.A.J. Metz
Horst R. Thieme

1 An introductory example: a model involving cannibalistic behaviour

Consider a population of, say, fish and assume that the behaviour of individuals (notably reproduction, survival, food consumption) depends on their *size* x. Let *birth size* be fixed at x_b and let *growth* be deterministic with a rate g that depends on size x and the x-specific *energy intake rate*. Likewise, *reproduction* is described by specifying how the rate β of producing offspring depends on size x and the x-specific energy intake rate. Assuming that predation is the only cause of death, we describe *survival* in terms of a death rate μ, which we call the x-specific *predation pressure*.

Next we have to describe how energy intake rate and predation pressure are themselves determined. For the purpose of exposition we assume that, apart from effects due to cannibalism, both food concentration and predation pressure are constant in time. In other words, cannibalism is the only *feedback loop* by which the individuals influence the environmental conditions of each other.

If we neglect the effects of handling times, satiation etc., we can describe the predation process in terms of attack rates and concentrations (the rate of eating being by definition the product, in this simple case of a linear functional response). Assume that individuals of size x have access to a food source with concentration $Z(x)$, which they attack at rate $C(x)$ and which has energetic value $E(x)$. Moreover, they attack conspecifics of size y with rate $c(x, y)$ and these have energetic value $e(y)$. Let m_t be the measure describing the population size and composition at time t (so $m_t(\omega)$ is the number (or rather spatial concentration) of individuals with size in ω at time t). Then the energy intake rate of an individual of size x is given by

$$I_1(t,x) = F_1(O_1(t,x),x) = E(x)C(x)Z(x) + O_1(t,x) \tag{1.1}$$

with

$$O_1(t,x) = \int_{[x_b,\infty)} e(y)c(x,y)m_t(dy) \tag{1.2}$$

while the predation pressure exerted on individuals of size x is given by

$$I_2(t,x) = F_2(O_2(t,x),x) = \sigma(x) + O_2(t,x) \tag{1.3}$$

with

$$O_2(t,x) = \int_{[x_b,\infty)} c(y,x)m_t(dy) \tag{1.4}$$

where σ denotes the x-specific non-cannibalistic predation pressure.

Once we now also specify g and β, we obtain a nonlinear structured population model, all interactions being due to cannibalism. We want to show constructively that one can associate in a meaningful way a dynamical system with such a model description. In other words, we want a construction that for given m_0 yields m_t, at least for t sufficiently small, and then we want to establish uniqueness in order to conclude that m_t qualifies as the population *state* (p-state) at time t.

To do so, we first cut the feedback loop and then re-establish it as a fixed point equation. More precisely, we pretend that the energy intake rate $I_1(t,x)$ and the predation pressure $I_2(t,x)$ are known functions. In this manner we obtain a non-autonomous (i.e., time dependent) but *linear* model to which we can apply the constructive procedure of Part I (Diekmann *et al.* 1998), culminating in solution operators that assign to a p-state m_0 the p-state m_t for $t > 0$ in a well-defined and unique manner. If we insert now these p-states in the formulas (1.2) and (1.4) we obtain a (nonlinear) input-output map. The biology expressed in (1.1) and (1.3) then requires that $I = F(O)$ which, since O depends on I, is a fixed point problem. Our aim in this paper is to show that for large classes of models one can derive Lipschitz estimates, apply the contraction mapping theorem to the fixed point problem to obtain a unique solution for any given initial p-state, and then use the fixed point to define a nonlinear autonomous dynamical system.

It is enticing to restrict the generality of c by putting

$$c(x,y) = \psi(x)\phi(y), \tag{1.5}$$

where ψ describes the degree of cannibalistic activity and ϕ the vulnerability to cannibalistic predation and where we have in mind that the support of ϕ is strictly to the left of the support of ψ (so big ones eat small ones but the precise size of potential predator and potential victim do influence what happens upon encounter in an independent manner; this is an example of making a model less parameter rich). The advantage is that we may now define outputs

$$\tilde{O}_1(t) = \int\limits_{[x_b,\infty)} e(y)\phi(y)m_t(dy) \tag{1.6}$$

and

$$\tilde{O}_2(t) = \int\limits_{[x_b,\infty)} \psi(y)m_t(dy), \tag{1.7}$$

which are only functions of time, and analyse how these depend on input when we take as the x-specific energy intake rate

$$I_1(t,x) = E(x)C(x)Z(x) + \psi(x)\tilde{O}_1(t) \tag{1.8}$$

and the x-specific predation pressure

$$I_2(t,x) = \sigma(x) + \phi(x)\tilde{O}_2(t). \tag{1.9}$$

However, to actually take advantage of this we must adapt the notion of input. We do so by noting that the x-specific energy intake rate is of the form

$$E(x)C(x)Z(x) + \psi(x)\tilde{I}_1(t) \tag{1.10}$$

and the x-specific predation pressure of the form

$$\sigma(x) + \phi(x)\tilde{I}_2(t) \tag{1.11}$$

and by now calling $\tilde{I}_1$ and $\tilde{I}_2$ the input. In this setting the feedback becomes simply the identity. We say that the environmental interaction variables are two-dimensional and we formulate the fixed point problem for $\mathbf{R}^2$-valued functions of time.

Although the framework we develop is of a more general nature, the class of examples we have in mind so far involves only $\mathbf{R}^k$-valued functions of time and we postpone a thorough analysis of inputs which are general functions of both t and x to some later time.

To reduce the parameter richness even further, we may choose

$$\psi(x) = \begin{cases} 0, & x < x_A \\ 1, & x \geq x_A \end{cases} \tag{1.12}$$

for some given x_A. This expresses that individuals become cannibalistic upon reaching size x_A (here A stands for "adult") and that there is no variation in the degree of cannibalistic tendency. The price we pay for such an "idealized" description is that individual behaviour changes abruptly as a function of individual state (i-state) x. When analysing the input-output-input map such a discontinuity needs special attention and in particular we need to make sure that the state of individuals always crosses the discontinuity transversally (in the present case of one dimensional i-state space this just means "with positive speed", but when the i-state space is higher dimensional the requirement is more easily interpreted as transversality (Diekmann *et al.* 2000)). In section 8 we return to this point.

The classical Holling time scale argument (see e.g. Metz and Diekmann (1986) and the references given therein) yields a saturating functional response reflecting a limited time budget and the effect of handling time. In the present situation involving size structure, we need to introduce a third interaction variable

$$I_3(t,x) = 1 + H(x)C(x)Z(x) + O_3(t,x), \tag{1.13}$$

with

$$O_3(t,x) = \int\limits_{[x_b,\infty)} h(x,y)c(x,y)m_t(dy), \tag{1.14}$$

where H and h are the respective handling times. The size-specific fraction of the time spent searching is then the inverse of $I_3(t,x)$. This fraction has to be incorporated in (1.1) and (1.3) as a multiplication factor, to account for the effect of handling time, thus letting us replace (1.1) and (1.4) by, respectively,

$$I_1(t,x) = \frac{E(x)C(x)Z(x) + O_1(t,x)}{I_3(t,x)} \tag{1.15}$$

and

$$O_2(t,x) = \int\limits_{[x_b,\infty)} \frac{c(y,x)}{I_3(t,y)} m_t(dy), \tag{1.16}$$

and thus introducing a dependence of the output on the instantaneous input while keeping the linearity of output in the p-state. We will show in Section 8 that a certain hierarchical structure makes the dependence of output on input rather harmless.

As a side-remark we mention that one can give another derivation of such expressions by invoking digestion as the limiting fast time scale process (Metz and Diekmann 1986).

$1\frac{1}{2}$ Trait d'union

We hope these modelling considerations have provided our readers with enough motivation, as well as enough understanding of the underlying general model structure, to dive into an abstract setting. Our approach will be top down. We start abstract and general and even trivial, in the sense that we simply assume everything we need. We work our way downwards by deriving in each step sufficient conditions for the assumptions in the preceding step. These may come in various forms and so we develop a theory with pyramid structure. The hope is that in this manner we may in the future incorporate new and essentially different examples with minimal effort, changing only the arguments in one (or a few) step(s).

2 Some terminology, definitions and hypotheses

Our basic thought experiment is that we

(i) pretend to know the state of the system at some initial time, which we take as the origin of the time axis;

(ii) pretend to know the input to the system for a length s of time;

(iii) determine the state of the system at time s.

Here an *input* is a function of time taking on values in a Banach space E. In the structured population context we call an element of E an “environmental condition” and the time argument of the input tells us at what time this condition is supposed to hold. An input I is defined on the interval $[0, \ell(I))$ and we call $\ell(I)$ the *length* of the input I.

It turns out to be convenient to introduce the *empty input* denoted by $\not I$. It is defined as an input of zero length: $\ell(\not I) = 0$. According to the definition of an input, it is an E-valued function defined on the empty interval $[0, 0)$ and thus it is nothing but the empty set. We have chosen the symbol $\not I$ which resembles the symbol $\emptyset$ for the empty set to remind us of this fact. The empty input $\not I$ should not be confused with the *zero input* 0, which is the function which has the constant value $0 \in E$ for all t in its interval of definition, which can have any length.

To inputs we can apply three basic operations, namely *restriction, shift* and *concatenation.* They are defined as follows:

Restriction ρ: For $0 \le s \le \ell(I)$, $\rho(s)I$ is the restriction of I to the subinterval $[0, s)$, that is,

$$(\rho(s)I)(t) = I(t) \quad \text{for} \quad 0 \le t < s.$$

Shift θ: The shift $\theta(-s)I$ is for $0 \le s \le \ell(I)$ defined on the interval $[0, \ell(I) - s)$ by

$$(\theta(-s)I)(t) = I(t + s),\ 0 \le t < \ell(I) - s.$$

Concatenation $\odot$: The concatenation $I_2 \odot I_1$ of I_1 and I_2 is defined on the interval $[0, \ell(I_1) + \ell(I_2))$ by

$$(I_2 \odot I_1)(t) = \begin{cases} I_1(t) & \text{for } 0 \le t < \ell(I_1), \\ I_2(t - \ell(I_1)) & \text{for } \ell(I_1) \le t < \ell(I_1) + \ell(I_2). \end{cases}$$

We collect some useful elementary properties into the following lemma:

Lemma 2.1

(i) $\rho(0)I = \not I$ *for all inputs* I*;*

(ii) $\rho(\ell(I))I = I$ *for all inputs* I*;*

(iii) $\rho(s)I = \rho(s)\rho(t+s)I$ *for all* $s \geq 0,\ t \geq 0$ *such that* $t + s \leq \ell(I)$*;*

(iv) $\theta(-0)I = I$ *for all inputs* I*;*

(v) $\theta(-\ell(I))I = \not{I}$ *for all inputs* I*;*

(vi) $\theta(-s)\theta(-t)I = \theta(-(s+t))I$ *for all* $s \geq 0,\ t \geq 0$ *such that* $t + s \leq \ell(I)$*;*

(vii) $\theta(-s)\rho(t+s)I = \rho(t)\theta(-s)I$ *for all* $s \geq 0,\ t \geq 0$ *such that* $t + s \leq \ell(I)$*;*

(viii) $I_3 \odot (I_2 \odot I_1) = (I_3 \odot I_2) \odot I_1$*;*

(ix) $I \odot \not{I} = \not{I} \odot I = I$ *for all inputs* I*;*

(x) $I = \theta(-s)I \odot \rho(s)I, \quad 0 \leq s \leq \ell(I)$.

There will be certain properties, like boundedness, measurability and integrability, that we require inputs to have. These properties should be such that they are preserved under restriction, shift and concatenation. Observe that, for instance, continuity is *not* preserved under concatenation. We also may want to identify inputs that differ only on sets of Lebesgue measure zero. As this identification commutes with the three basic operations, they extend to equivalence classes that are obtained by the identification.

To formalize the setting, we have to postulate certain properties of the spaces to which the inputs belong. Because the inputs may have arbitrary lengths we have to introduce a whole family $\{B_s\}_{s\geq 0}$ of spaces. Here and in the following hypothesis the parameter s should be interpreted as the length of an input.

Hypothesis 2.2

(a) $B_0 = \{\not{I}\}$ and for a given but arbitrary $s > 0$, B_s is a set of (equivalence classes of) functions defined on the interval $[0, s)$ with values in E such that

  (1) for $0 \leq \sigma \leq s$ the restriction $\rho(\sigma)$ maps B_s onto B_σ,

  (2) for $0 \leq \sigma \leq s$ the shift $\theta(-\sigma)$ maps B_s onto $B_{s-\sigma}$,

  (3) for $s_1 \geq 0,\ s_2 \geq 0$ concatenation is a one-to-one mapping of $B_{s_1} \times B_{s_2}$ onto $B_{s_1+s_2}$,

  (4) the constant functions defined on $[0, s)$ belong to B_s.

(b) For each $s \geq 0$, B_s is a Banach space with norm $||\cdot||$ (note that the norm depends on s but that we do not express this in the notation) such that

  (1) for $0 < \sigma < s$, $\rho(\sigma)$ and $\theta(-\sigma)$ are bounded linear operators of norm one (the same is true for $\rho(s)$ and $\theta(0)$, but $\rho(0)$ and $\theta(-s)$ have norm zero),

  (2) $||I \odot 0|| = ||I|| = ||0 \odot I||$

  (3) $I_2 \odot 0 + 0 \odot I_1 = I_2 \odot I_1$

In assertion (b3) above, the lengths of the zero inputs are of course assumed to be such that the sum makes sense, that is, such that both terms on the left hand side have the same length. It follows from (b2) and (b3) that

$$||I_2 \odot I_1|| \leq ||I_2|| + ||I_1||. \tag{2.1}$$

Despite a slight abuse of the symbol ρ, it seems natural to denote the input defined on $[0, s)$ taking the constant value $\overline{I} \in E$ by $\rho(s)\overline{I}$. With this convention assertion (a4) can be written as $\rho(s)\overline{I} \in B_s$.

We use $\mathcal{B}$ to denote $\bigcup_{s\geq 0} B_s$. Note that for $I \in \mathcal{B}$ we have $\ell(I) = s$ if and only if $I \in B_s$. Moreover, it follows from Lemma 2.1 (viii) and (ix) that $\mathcal{B}$ is a monoid (that is, a semigroup with a unit element) under concatenation, with the empty input $\not{I}$ as unit.

In Hypotheses 2.2 (a) we formalized the requirement that certain technical constraints on the inputs are preserved under restriction, shift, and concatenation. But usually the biological interpretation also puts constraints on the inputs and in most cases these take the form of a condition on the range of the inputs. A typical example is when the interpretation requires the input to take on only nonnegative values. We shall therefore from now onassume that the inputs take on values in a subset Z of the Banach space E. Obviously this range condition is invariant under the three basic operations.

We denote the subset of B_s consisting of functions with values in Z by $B_s(Z)$. Likewise we use $\mathcal{B}(Z)$ to denote $\bigcup_{s\geq 0} B_s(Z)$.

Let Y be a set. The set Y figures as the state space of the dynamical system that we want to construct. We now formulate the assumption that for a *given* input we have a well-defined dynamical system.

Hypothesis 2.3 *(The semigroup property)* For every $I \in \mathcal{B}(Z)$ there exists a map T_I from Y to Y such that

$$T_{\not{I}} = \mathrm{id}_Y, \tag{2.2}$$

$$T_{I_2}\, T_{I_1} = T_{I_2 \odot I_1}. \tag{2.3}$$

In (2.2) id_Y is the identity mapping on Y. Note that (2.3) can equivalently be stated as

$$T_I = T_{\theta(-\sigma)I}\, T_{\rho(\sigma)I}, \quad 0 \leq \sigma \leq \ell(I). \tag{2.4}$$

Note that the information about how much we go forward in time is contained in the length of the input. Whenever there is a need to consider, for given I, the population states for times t with $0 \leq t < \ell(I)$, we do so by means of the restriction operator, that is, by considering $T_{\rho(t)I}y$.

The name "semigroup" derives from the fact that (2.3) states that the map $I \mapsto T_I$ from $\mathcal{B}(Z)$ to the set of maps of Y into itself (which is a semigroup under composition) is a semigroup homomorphism. As a matter of fact it is even a monoid homomorphism as (2.2) says that the unit of $\mathcal{B}(Z)$ is mapped to id_Y.

For constant inputs we obtain semigroups of maps of Y into Y parametrized by positive real numbers. Indeed, for $\overline{I} \in Z$, define $\overline{T}(s) = T_{\rho(s)\overline{I}}$. Then

$$\overline{T}(s_1)\overline{T}(s_2) = T_{\rho(s_1)\overline{I}}T_{\rho(s_2)\overline{I}} = T_{\rho(s_1)\overline{I}\odot\rho(s_2)\overline{I}} = T_{\rho(s_1+s_2)\overline{I}} = \overline{T}(s_1+s_2). \tag{2.5}$$

3 Construction of a dynamical system (closing the feedback loop)

To define the output, we introduce a map $H : Y \to Z$. In the setting of Hypothesis 2.3, let $y \in Y$ and $I \in \mathcal{B}(Z)$ be given. The *output* is then the function

$$t \mapsto H\left(T_{\rho(t)I}\; y\right) \tag{3.1}$$

defined on $[0, \ell(I))$ and with values in Z. We are here, for the sake of mathematical simplicity, thinking of a feedback map which is the identity, such that the distinction between the input-output-input map and the input-output map introduced in Definition 3.2 below becomes irrelevant. Relative to the formulation which corresponds most closely to the biological mechanism this may entail a mathematical transformation, as, e.g., the step from $I(t,x)$ to $\tilde{I}(t)$ in Section 1.

Hypothesis 3.1 The output defined by (3.1) is an element of $B_{\ell(I)}(Z)$.

Next we introduce the map that transforms input into output, given the population state y.

Definition 3.2 For each $y \in Y$ the *input-output map* $P_y : \mathcal{B}(Z) \to \mathcal{B}(Z)$ is defined by

$$P_y(I) = H\left(T_{\rho(\cdot)I}\ y\right).$$

By Hypothesis 3.1 P_y maps $B_s(Z)$ into $B_s(Z)$ for each $s \geq 0$. Moreover, P_y commutes with $\rho(\sigma)$ for all $\sigma \geq 0$.

A shift in the input should be reflected in a corresponding shift in the output, provided the population state is updated accordingly. That this is indeed the case is shown in the following lemma.

Lemma 3.3 *For all* $I \in \mathcal{B}(Z)$ *and all* $0 \leq s \leq \ell(I)$ *one has*

$$\theta(-s)P_y(I) = P_{T_{\rho(s)I}\ y}\left(\theta(-s)I\right). \tag{3.2}$$

Proof. If $s = \ell(I)$, then (3.2) reduces to the identity $\not{I} = \not{I}$. For $0 \leq s < \ell(I)$, $0 \leq t < \ell(I) - s$ the left hand side of (3.2) evaluated at t equals $H\left(T_{\rho(t+s)I}\ y\right)$ whereas the right hand side equals $H\left(T_{\rho(t)\theta(-s)I}\ T_{\rho(s)I}\ y\right)$. It follows from Lemma 2.1 and the semigroup property of Hypothesis 2.3 that the two sides are indeed equal. □

The distinction between input and output is, in our context, a mental construction and the two should in fact be identical. In other words, our task is to find a fixed point of the map P_y for arbitrary $y \in Y$. At this level in our top down approach we state this as a hypothesis:

Hypothesis 3.4 For all $y \in Y$ there exists an $s(y) > 0$ such that $P_y\ |_{B_s(Z)}$ has a *unique* fixed point, to be called I_y, for every $s \leq s(y)$.

Strictly speaking the fixed point not only depends on y but also on the s that we choose. However, the fixed point on a smaller interval is simply the restriction of the fixed point on a larger interval (because of uniqueness and the fact that the restriction map commutes with P_y) and therefore we may safely suppress s in the composite symbol denoting the fixed point.

Lemma 3.5 *For all* $y \in Y$, $s \in [0, s(y))$ *one has*

$$\theta(-s)I_y = I_{T_{\rho(s)I_y}\ y}.$$

Proof. One has $\theta(-s)I_y = \theta(-s)P_y\left(I_y\right) = P_{T_{\rho(s)I_y}\ y}\left(\theta(-s)I_y\right)$ by Lemma 3.3. So $\theta(-s)I_y$ is a fixed point of $P_{T_{\rho(s)I_y}\ y}$ and by uniqueness it must therefore be equal to $I_{T_{\rho(s)I_y}\ y}$. □

Definition 3.6 For $t \geq 0$ we put

$$S(t, y) = T_{\rho(t)I_y}\ y \tag{3.3}$$

whenever the right hand side is defined.

Note that it follows from (3.3) that

$$S(0, y) = T_{\rho(0)I_y}\ y = T_{\not{I}}\ y = \mathrm{id}_Y y = y \tag{3.4}$$

for all $y \in Y$.

Theorem 3.7 *Let $y \in Y$. Then $s < s(y)$, $t < s(S(s,y))$ implies $t + s < s(y)$ and*

$$S(t+s,y) = S(t,S(s,y)). \tag{3.5}$$

Proof.

$$\begin{aligned} S(t,S(s,y)) &= T_{\rho(t)I_{S(s,y)}}\ S(s,y) = T_{\rho(t)I_{T_{\rho(s)I_y}\ y}}\ S(s,y) \\ &= T_{\rho(t)\theta(-s)I_y}\ S(s,y) = T_{\theta(-s)\rho(t+s)I_y}\ T_{\rho(s)I_y}\ y \\ &= T_{\theta(-s)\rho(t+s)I_y \odot \rho(s)I_y}\ y = T_{\rho(t+s)I_y}\ y \\ &= S(t+s,y). \end{aligned}$$

□

Theorem 3.7 together with the identity (3.4) says that S is a semiflow. Usually one requires that a semiflow is continuous both with respect to time and initial state.

Whenever we verify Hypothesis 3.4 we say that we have *solved* a nonlinear problem, meaning, of course, that we can combine T_I and I_y into a semiflow via (3.3)

4 Kernels and convolutions

As in Part I (Diekmann *et al.* 1998) we consider individual states as elements of a measurable space Ω with a countably generated σ-algebra Σ. Our use of the word "kernel" is somewhat different from that of Part I. Here a *kernel* k is a map from $\Omega \times \Sigma$ into $\mathbf{R}$ which is bounded and measurable with respect to the first variable and countably additive with respect to the second variable. (So for fixed $\omega \in \Sigma$ the function $x \mapsto k(x,\omega)$ is bounded and measurable, while for fixed $x \in \Omega$ the map $\omega \mapsto k(x,\omega)$ defines a finite signed measure on Ω). We call a kernel *positive* if it assumes non-negative values only.

The *product* $k^1 \times k^2$ of two kernels k^1 and k^2 is the kernel defined by

$$\left(k^1 \times k^2\right)(x,\omega) = \int_\Omega k^1(\xi,\omega)k^2(x,d\xi). \tag{4.1}$$

Likewise we define the *product* $f \times k$ of a bounded measurable function $f : \Omega \to Z$ and a kernel k as the function

$$(f \times k)(x) = \int_\Omega f(\xi)k(x,d\xi). \tag{4.2}$$

The product of a kernel k and a measure μ is defined analogously as the measure

$$(k \times \mu)(\omega) = \int_\Omega k(\xi,\omega)\mu(d\xi). \tag{4.3}$$

Finally we agree that the product $f \times \mu$ of a function and a measure is

$$f \times \mu = \int_\Omega f(x)\mu(dx) \in Z. \tag{4.4}$$

The $\times$-product is associative in the following sense: If in the case of three objects f, k and μ, say, both the products $(f \times k) \times \mu$ and $f \times (k \times \mu)$ are well-defined, then they are equal. In this case we leave out the parentheses and write simply $f \times k \times \mu$.

We shall use inputs $I \in \mathcal{B}$ to parametrize kernels and functions. For two parametrized families k^1_I and k^2_I of kernels we define their *convolution product* $\left(k^1 * k^2\right)_I$ by

$$\left(k^1 * k^2\right)_I = \int_{[0,\ell(I))} k^1_{\theta(-\sigma)I} \times k^2_{\rho(d\sigma)I}, \tag{4.5}$$

whenever the integral exists. In particular, this is the case if $\sigma \mapsto k^1_{\theta(-\sigma)I}(\xi,\omega)$ is bounded, uniformly in ξ and ω, and measurable, while $\sigma \mapsto k^2_{\rho(\sigma)I}(x,\omega)$ is of bounded variation uniformly in x and ω.

The convolution product of a parametrized family f_I of functions and a parametrized family k_I of kernels is defined analogously:

$$(f * k)_I = \int_{[0,\ell(I))} f_{\theta(-\sigma)I} \times k_{\rho(d\sigma)I}. \tag{4.6}$$

Note that the convolution of two parametrized families of kernels is again a parametrized family of kernels, while the convolution of a family of functions and kernels yields a family of functions.

When deriving Lipschitz estimates for the input-output map P_y, we need sup-norm estimates for convolution products. To prepare the way, first note that

$$|(f \times k)\,(x)|_E \leq \sup_{\xi\in\Omega} |f(\xi)|_E\, |k|(x,\Omega), \tag{4.7}$$

where $|\cdot|_E$ denotes the norm in the Banach space E, $|k|(x,\cdot)$ denotes the total variation measure of $k(x,\cdot)$ and accordingly $|k|(x,\Omega)$ is the total variation of $k(x,\cdot)$. We also need the total variation of a real valued function ϕ defined on an interval $[0,s)$. This will be denoted by $V(\phi)$. As the length of the interval will always be clear from the context it need not be included in the symbol.

We now lift the inequality (4.7) to the convolution product. The rationale for the introduction of the subset Ω_b of Ω will be explained in the next section.

Lemma 4.1 *Let f_I and k_I be parametrized families of functions and kernels, respectively. If $k(x,\cdot)$ is concentrated on Ω_b for all $x \in \Omega$, then*

$$|(f * k)_I\,(x)|_E \leq \sup_{\xi\in\Omega_b,\, \sigma\in[0,\ell(I))} \left|f_{\theta(-\sigma)I}(\xi)\right|_E V\left(\left|k_{\rho(\cdot)I}\right|(x,\Omega_b)\right). \tag{4.8}$$

Proof. One has

$$(f * k)_I = \int_{[0,\ell(I))} f_{\theta(-\sigma)I} \times k_{\rho(d\sigma)I}$$

and hence, by (4.7)

$$|(f * k)_I\,(x)|_E \leq \int_{[0,\ell(I))} \sup_{\xi\in\Omega_b} \left|f_{\theta(-\sigma)I}(\xi)\right| \left|k_{\rho(d\sigma)I}\right|(x,\Omega_b).$$

From here the inequality (4.8) follows directly. ◻

5 Linear structured population models with input

When modelling structured populations one starts by describing individual behaviour. A first task of the mathematician is then to show that this description leads to a well-defined dynamical system at the population level, that is, a dynamical system that for any given initial population state gives the population state for future instants of time.

Usually individual behaviour is described in terms of rates of development, death and reproduction. In (Diekmann *et al.* 1998) we argued at length that a certain pre-processing of such basic ingredients, leading to composite ingredients at a somewhat higher level of aggregation, has conceptual and technical advantages. In this section we take this pre-processing step for granted, but in Section 8 we return to this point.

Let Ω be a measurable space with a countably generated σ-algebra Σ. Individuals are characterized by their i-state, which is represented by an element x of Ω. Ω is therefore called the i-state space. The two *ingredients* of a *linear* structured population model with input are two parametrized families u_I and Λ_I of kernels which have the following interpretations:

- $u_I(x, \omega)$ is the probability that, given the input I, an individual which has i-state $x \in \Omega$ at a certain time, is still alive $\ell(I)$ time units later and then has i-state in $\omega \in \Sigma$.
- $\Lambda_I(x, \omega)$ is the expected number of offspring, with state-at-birth in $\omega \in \Sigma$, produced by an individual, with i-state $x \in \Omega$ at a certain time, within the time interval of length $\ell(I)$ following that time, given the input I.

The interpretation of the ingredients u_I and Λ_I requires that certain consistency relations and monotonicity conditions hold. We collect these into the following assumption:

Assumption 5.1

(i) u_I and Λ_I are parameterized families of positive kernels.

(ii) For every I_1 and I_2 in $\mathcal{B}(Z)$ one has

$$u_{I_2 \odot I_1} = u_{I_2} \times u_{I_1}.$$

(iii) For every I_1 and I_2 in $\mathcal{B}(Z)$ one has

$$\Lambda_{I_2 \odot I_1} = \Lambda_{I_1} + \Lambda_{I_2} \times u_{I_1}.$$

(iv) For any $x \in \Omega$, $\omega \in \Sigma$, $I \in \mathcal{B}(Z)$ the function $\sigma \mapsto \Lambda_{\rho(\sigma)I}(x, \omega)$ is non-decreasing and

$$\lim_{\sigma \downarrow 0} \Lambda_{\rho(\sigma)I}(x, \omega) = \Lambda_{\not{I}}(x, \omega) = 0.$$

(v) For any $x \in \Omega$, $\omega \in \Sigma$, $I \in \mathcal{B}$ the function $\sigma \mapsto u_{\rho(\sigma)I}(x, \Omega)$ is non-increasing and

$$\lim_{\sigma \downarrow 0} u_{\rho(\sigma)I}(x, \omega) = u_{\not{I}}(x, \omega) = \delta_x(\omega),$$

In particular,

$$u_I(x, \Omega) \leq 1.$$

Relation (ii) is nothing but the Chapman–Kolmogorov equation, while relation (iii) is a similar consistency relation tying reproduction, survival and individual development together (see Diekmann *et al.* 1998 for more motivation). Sometimes we require in addition to (v) that

$$\lim_{\ell(I) \to \infty} u_I(x, \Omega) = 0 \tag{5.1}$$

uniformly for $x \in \Omega$ or the somewhat stronger condition of a uniformly bounded life expectancy: There exists an $M < \infty$ such that

$$\int_{[0, \ell(I))} u_{\rho(\sigma)I}(x, \Omega) d\sigma \leq M \tag{5.2}$$

for every $x \in \Omega$ and every $I \in \mathcal{B}(Z)$. The limit in (5.1) is of the general type $\lim_{\ell(I) \to \infty} f_I = g$, which in an arbitrary metric space is defined by $\forall \varepsilon > 0\ \exists M > 0$ such that $I \in \mathcal{B}(Z)$, $\ell(I) > M \Rightarrow d(f_I, g) < \varepsilon$.

The population state (p-state) is by definition the distribution of i-states and can therefore be represented by a measure m on the i-state space Ω. A natural choice for the p-state space is therefore a closed subcone Y of $\mathrm{M}_+(\Omega)$, the cone of all (finite) positive measures on Ω. The dynamical system T_I describing the dynamics at the population level should therefore be such that given the initial p-state m_0 and the input I, $T_I\ m_0$ is the p-state at time $\ell(I)$. The population at time $\ell(I)$ consists of those individuals present in the initial population that are still alive and all living descendants of the initial population. Suppose that we have somehow been able to construct a measure $u_I^c(x,\cdot)$ on Ω with the interpretation that $u_I^c(x,\omega)$ is $u_I(x,\omega)$ plus the expected number of descendants (i.e. children, grand-children, great grand-children, etc) of an individual initially of i-state x, which are still alive and have i-state in ω, $\ell(I)$ time units later. Here the superscript c refers to "clan". Summing up over all individuals present initially we obtain the composition of the population at time $\ell(I)$ as follows:

$$(T_I\ m_0)(\omega) = \int_\Omega u_I^c(x,\omega) m_0(dx). \tag{5.3}$$

Suppose furthermore that we have constructed Λ_I^c with the same interpretation as Λ_I, but now referring to the whole clan. Because every member of the clan is either a child of the ancestor or a child of a member of the clan, or alternatively, either a child of the ancestor or a member of the clan of a child of the ancestor, we obtain the following consistency relation:

$$\Lambda_I^c = \Lambda_I + (\Lambda * \Lambda^c)_I = \Lambda_I + (\Lambda^c * \Lambda)_I\,. \tag{5.4}$$

We now notice that we only have to construct Λ_I^c because once this has been done, the verbal description of u_I^c can be formalized as

$$u_I^c = u_I + (u * \Lambda^c)_I\,. \tag{5.5}$$

The general linear structured population problem with input can now be formulated as follows:

Linear structured population problem with input. Given the ingredients u_I and Λ_I, construct Λ_I^c such that (5.4) holds for every $I \in \mathcal{B}(Z)$ and show that the family $\{T_I\}_{I\in\mathcal{B}(Z)}$ of linear operators on (the span of) Y defined by (5.3) and (5.5) is a semigroup.

The state-at-birth is really a *state*, that is, it summarizes all information that is relevant for predicting the future. Hence the expected number of grand-children is obtained as the convolution product of Λ_I with itself, the expected number of great-grand-children as the threefold convolution product of Λ_I with itself, etc.. The clan is obtained by summing up over all generations:

$$\Lambda_I^c = \sum_{k=1}^{\infty} \Lambda_I^{k*}, \tag{5.6}$$

In (5.6) $\Lambda_I^{1*} = \Lambda_I$ and $\Lambda_I^{k*} = \left(\Lambda^{(k-1)*} * \Lambda\right)_I$ for $k \geq 2$.

The positivity of the family Λ_I guarantees that (5.6) has a meaning in any case, but additional conditions on Λ_I (e.g. a reproduction delay preventing newborns to give birth) guarantee that the sum converges to something finite (Diekmann *et al.* 1998). Another important feature that often simplifies the analysis is that Λ_I may be concentrated on a set $\Omega_b \subset \Omega$ which may be considerably smaller that Ω itself (indeed, Ω_b may consist of just one point: $\Omega_b = \{x_b\}$ as in the introductory example of Section 1). We formalize these ideas in the following definition (cf. Diekmann *et al.* 1998, Definitions 2.5 and 2.7).

Definition 5.2 (i) A set $\Omega_b \in \Sigma$ is called a *set representing the birth states* if the measure $\Lambda_I(x,\cdot)$ is concentrated on Ω_b for all $x \in \Omega$ and all $I \in \mathcal{B}(Z)$.

(ii) $x \in \Omega$ is called a *state with reproduction delay at least* ε if $\Lambda_I(x, \Omega) = 0$ for all $I \in B_s(Z)$ with $s < \varepsilon$.

Obviously we would like to choose Ω_b as small as possible. But as we already pointed out in (Diekmann *et al.* 1998) there is, in general, no unique way of achieving this goal. To see this, notice that if Ω_b is a set representing the birth states one can remove from Ω_b any set ω such that $\Lambda_I(x, \omega) = 0$ for all $x \in \Omega$ and all $I \in \mathcal{B}(Z)$ without destroying property (i) of Definition 5.2. But one can certainly not remove an uncountable union of such sets. If Ω has a natural topology, then one can use the idea of *support* of a measure and define Ω_b to be the smallest *closed* set such that $\Lambda_I(x, \neg\Omega_b) = 0$ for all $x \in \Omega$ and all $I \in \mathcal{B}(Z)$ (here $\neg$ denotes the complement of a set).

The interpretation of u_I^c and Λ_I^c given above requires that u_I^c and Λ_I^c, too, satisfy the Chapman-Kolmogorov equation and the reproduction-survival-i-state-development consistency relation. That this is indeed the case was proved (in a slightly different setting) in Part I (Diekmann *et al.* 1998), where we also showed that Λ_I^c is the (unique) *resolvent* of Λ_I. We collect these facts into a proposition:

Proposition 5.3

(i) *For every* I_1 *and* I_2 *in* $\mathcal{B}(Z)$ *one has*

$$u^c_{I_2 \odot I_1} = u^c_{I_2} \times u^c_{I_1}.$$

(ii) *For every* I_1 *and* I_2 *in* $\mathcal{B}(Z)$ *one has*

$$\Lambda^c_{I_2 \odot I_1} = \Lambda^c_{I_1} + \Lambda^c_{I_2} \times u^c_{I_1}.$$

(iii) Λ_I^c *defined by* (5.6) *is the unique solution of Equation* (5.4)*and*

$$u_I^c = u_I + (u^c * \Lambda)_I\,. \tag{5.7}$$

The map $T_I : Y \to Y$ is now defined by $T_I\ m_0 = u_I^c \times m_0$, that is, by (5.3). By Proposition 5.3 (i) T_I is indeed a semigroup, that is, it satisfies Hypothesis 2.3. By the uniqueness result (iii) of Proposition 5.3 this is the only semigroup describing the dynamics at the population level. We can thus summarize the contents of (Diekmann *et al.* 1998) as follows:

Theorem 5.4 *Under Assumption* 5.1*, the linear structured population problem with input has a unique solution.*

6 Nonlinear structured population models

In the previous section we showed that under Assumption 5.1 the model ingredients u_I and Λ_I uniquely determine a linear semigroup $\{T_I\}_{I \in \mathcal{B}(Z)}$ on (the span of) the p-state space. In this section we shall formulate nonlinear population problems, where the input I is not given beforehand but fed back into the system from an output.

When the output is obtained by applying a *linear* map from the p-state space $Y \subset \mathrm{M}_+(\Omega)$ to Z we speak about a *pure mass action problem*. In this case we shall actually assume slightly more, viz. that the output map $H : Y \to Z$ is represented by

$$H(m) = \gamma \times m = \int_\Omega \gamma(x) m(dx) \tag{6.1}$$

for some *bounded* and measurable $\gamma : \Omega \to Z$. So, in the pure mass action case the specification of the nonlinear problem requires only one new ingredient: γ.

The $\times$-product allows us to give a nice representation of the input-output-input map P_y:

$$P_{m_0}(I) = \gamma \times u^c_{\rho(\cdot)I} \times m_0. \tag{6.2}$$

We are now ready to formulate our first nonlinear structured population problem.

Pure mass action problem. Given the ingredients u_I, Λ_I and γ and the initial p-state $m_0 \in Y$, show that the input-output-input map P_{m_0} defined by (6.2) has a unique fixed point I_{m_0} in some space $B_s(Z)$. The dynamical system describing the time-evolution of the p-state is then given by

$$S(t, m_0) = T_{\rho(t)I_{m_0}} m_0$$

and we say that the problem has been solved.

Remark 6.1 Preferably there should be a uniform (in m_0) lower bound for $s = s(m_0)$, since such a bound guarantees global existence. As the reasons for the existence of such a bound (and hence the techniques for deriving the bound) are quite problem specific, we do not deal with the issue in the current paper. Whenever a model is based on energy budget considerations (Kooijman 2000) we expect that global existence is guaranteed.

To solve the pure mass action problem one has to verify that P_{m_0} maps $B_s(Z)$ into $B_s(Z)$ and is, for s sufficiently (depending on m_0) small, a contraction mapping once $B_s(Z)$ is equipped with a suitable norm.

As we have seen in Section 1, time scale arguments applied to mass action model formulations may lead to more complicated outputs, which either can be represented by a nonlinear map H on Y or by a linear map on Y depending on the input itself. In such cases there seems to be (always, as far as we know) a *hierarchical structure* in the sense that $Z = Z_1 \times Z_2 \times \cdots \times Z_k$ and, in self explaining notation,

$$\begin{aligned} H_1(y) &= L_1(y), \\ H_2(y) &= L_2(I_1, y) = L_2(L_1(y), y), \\ &\vdots \\ H_k(y) &= L_k(I_1, I_2, \ldots, I_{k-1}, y). \end{aligned} \tag{6.3}$$

We call the resulting nonlinear structured population problem a **generalized mass action problem.** We shall concentrate on the case $k = 2$ and formulate our results in such a way that an induction argument settles the case of a general k.

For the time being, let us restrict ourselves to the case of linear output. Define, for $i = 1$ or c,

$$o^i_I = \gamma \times u^i_I. \tag{6.4}$$

Multiplying (in the sense of $\times$) equation (5.7) from the left by γ one obtains

$$o^c_I = o_I + (o^c * \Lambda)_I \tag{6.5}$$

and this is the equation we are going to analyse in the next section. What we shall do is formulate assumptions on the maps $I \mapsto o_I$ and $I \mapsto \Lambda_I$ and derive conclusions about the map $I \mapsto o^c_I$ which take the form of a Lipschitz estimate with a constant that tends to zero as $\ell(I) \downarrow 0$.

7 Lipschitz estimates

In our top down spirit we now start working downwards to derive sufficient conditions for the assumptions concerning P_y to hold. We start by a lemma.

Lemma 7.1 *Suppose there exists a bounded and measurable function* $C_1 : \Omega \times \mathbf{R}_+ \to \mathbf{R}_+$ *such that for* o_I^c *defined by* (6.4) *one has*

$$\left|\left|o^c_{\rho(\cdot)I}(x) - o^c_{\rho(\cdot)J}(x)\right|\right| \leq C_1(x,s)\,||I-J||\,, \quad x \in \Omega, \tag{7.1}$$

for all I *and* J *in* $B_s(Z)$ *and such that*

$$C_1(x,s) \downarrow 0 \tag{7.2}$$

for all $x \in \Omega$. *Then Hypothesis* 3.4 *is verified for* $Y = M_+(\Omega)$ *and* P *defined by* (6.2).

The idea of the proof of Lemma 7.1 is simple. Evidently (7.2) implies that

$$\lim_{s \downarrow 0} \int_\Omega C_1(x,s) m_0(dx) = 0 \tag{7.3}$$

for all $m_0 \in M_+(\Omega)$. Then, because by (6.2) and (6.4) one has

$$P_{m_0}(I) = o^c_{\rho(\cdot)I} \times m_0, \tag{7.4}$$

we can apply the contraction mapping principle and conclude that Hypothesis 3.4 is indeed satisfied.

In this section we shall provide assumptions on Λ_I and o_I, which together with (6.5) imply the estimate (7.1) and hence yield existence and uniqueness of solutions of the population problem.

So far the spaces $B_s(Z)$ and, in particular, the norm on them, have not been specified. But gradually we need to become more specific. In the remainder of this section the norm is either the L_1-norm $\|\cdot\|_1$, in which case $B_s(Z) = \{I \in L_1([0,s);E) : I(t) \in Z \text{ for almost all } t\}$ or the sup-norm $\|\cdot\|_\infty$, in which case $B_s(Z)$ is either the space of bounded measurable functions on $[0,s)$ with values in Z or the space of regulated functions with this domain and range (we define regulated functions as the uniform limits of step functions, see (Dieudonné 1969, p. 145)). To understand why we restrict ourselves to these choices, recall that we need invariance under concatenation.

When deriving estimates below, the following lemma will come in helpful. We start by introducing some notation.

Let $\Omega_b \subset \Omega$. For a function $f : \mathcal{B}(Z) \times \Omega \to \mathbf{R}$ we define $\overline{f} : \mathcal{B}(Z) \to \mathbf{R}_+$ by

$$\overline{f}(I) = \sup_{\xi \in \Omega_b,\, \alpha \in [0,\ell(I))} |f(\theta(-\alpha)I, \xi)|\,, \quad I \in \mathcal{B}(Z). \tag{7.5}$$

When the argument $I \in \mathcal{B}(Z)$ of f is written as a subscript, the same convention is used for $\overline{f}$. Thus, for instance, we write $\overline{o}^c_I$. If f has a third argument, which is kept constant when taking the supremum in (7.5), then $\overline{f}$ gets an additional argument: for instance

$$\overline{\Lambda}_I(\Omega_b) = \sup_{\xi \in \Omega_b,\, \alpha \in [0,\ell(I))} \Lambda_{\theta(-\alpha)I}(\xi, \Omega_b).$$

Lemma 7.2 *Let ϕ, h and K be functions defined on $\mathcal{B}(Z) \times \Omega$ with values in $\mathbf{R}_+$ and assume that*

$$\phi(I,x) \leq h(I,x) + \overline{\phi}(I) K(I,x) \tag{7.6}$$

and

$$\overline{K}(I) < 1. \tag{7.7}$$

Then, for $\ell(I)$ sufficiently small,

$$\phi(I,x) \leq h(I,x) + \left(1 - \overline{K}(I)\right)^{-1} \overline{h}(I) K(I,x). \tag{7.8}$$

Proof. Replacing I by $\theta(-\alpha)I$ in (7.6) and noting that $\overline{\phi}(\theta(-\alpha)I) \leq \overline{\phi}(I)$ we obtain

$$\phi(\theta(-\alpha)I, x) \leq h(\theta(-\alpha)I, x) + \overline{\phi}(I) K(\theta(-\alpha)I, x).$$

So taking the supremum over $x \in \Omega_b$ and $\alpha \in [0, \ell(I))$ we find

$$\overline{\phi}(I) \leq \overline{h}(I) + \overline{\phi}(I)\overline{K}(I),$$

which, under the assumption (7.7), implies

$$\overline{\phi}(I) \leq \left(1 - \overline{K}(I)\right)^{-1} \overline{h}(I). \tag{7.9}$$

Inserting (7.9) into (7.6) we find (7.8). ◻

Our first estimate gives a bound on o_I^c in terms of bounds on o_I and Λ_I.

Lemma 7.3 *Assume that there are positive constants K_1 and K_2 and a nondecreasing function $C_1(s)$, with $\lim_{s\downarrow 0} C_1(s) = 0$, such that one has for all $x \in \Omega$ and all $I \in \mathcal{B}(Z)$*

$$|o_I(x)|_E \leq K_1 \tag{7.10}$$

$$\Lambda_I(x, \Omega_b) \leq K_2 \tag{7.11}$$

and for all $x \in \Omega_b$

$$\Lambda_I(x, \Omega_b) \leq C_1(\ell(I)). \tag{7.12}$$

Then, provided $C_1(\ell(I)) < 1$,

$$|o_I^c(x)|_E \leq K_1 \left(1 + (1 - C_1(\ell(I)))^{-1} K_2\right). \tag{7.13}$$

Proof. If we take the E-norm of both sides of the convolution equation (6.5) we find by virtue of Lemma 4.1 the inequality (7.6) with

$$\begin{aligned} \phi(I,x) &= |o_I^c(x)|_E \\ h(I,x) &= |o_I(x)|_E \\ K(I,x) &= \Lambda_I(x, \Omega_b) \end{aligned}$$

and so the conclusion of Lemma 7.2 yields the estimate

$$|o_I^c(x)|_E \leq |o_I(x)|_E + \left(1 - \overline{\Lambda}_I(\Omega_b)\right)^{-1} \overline{o}_I \Lambda_I(x, \Omega_b) \tag{7.14}$$

provided (7.7) is true. Note that (7.12) guarantees that (7.7) is true for $\ell(I)$ sufficiently small. Inserting the bounds (7.10), (7.11), and (7.12) into the estimate (7.14) we arrive at (7.13). ◻

When estimating differences of outputs we are led to consider *doubly* parametrized families. For inputs I_1 and I_2 of equal length (i.e., $\ell(I_1) = \ell(I_2)$) we define the $Z \times Z$ valued function I by $I = (I_1, I_2)$. We then define the convolution product as before; cf. (4.5) and (4.6). Starting from the two equations

$$o_I^{ic} = o_I^i + \left(o^{ic} * \Lambda^i\right)_I, \quad i = 1, 2,$$

where $o_I^i = o_{I_i}$, etc., we arrive by subtraction and rearrangement at

$$o_I^{1c} - o_I^{2c} = o_I^1 - o_I^2 + \left(o^{1c} * \left(\Lambda^1 - \Lambda^2\right)\right)_I + \left(\left(o^{1c} - o^{2c}\right) * \Lambda^2\right)_I$$

or

$$o_I^{1c} - o_I^{2c} = g_I + \left(\left(o^{1c} - o^{2c}\right) * \Lambda^2\right)_I \tag{7.15}$$

with

$$g_I := o_I^1 - o_I^2 + \left(o^{1c} * \left(\Lambda^1 - \Lambda^2\right)\right)_I. \tag{7.16}$$

So the difference $o_{I_1}^c - o_{I_2}^c$ satisfies a convolution equation with forcing function g_I. We proceed by deriving an estimate for g_I.

Lemma 7.4 *One has*

$$|g_I(x)|_E \leq |o_{I_1}(x) - o_{I_2}(x)|_E + \overline{o_{I_1}^c} V\left(\left|\Lambda_{\rho(\cdot)I_1} - \Lambda_{\rho(\cdot)I_2}\right|(x, \Omega_b)\right). \tag{7.17}$$

Proof. This is nothing but Lemma 4.1 applied to the particular situation. ◻

Lemma 7.5 *Assume* (7.11) *and* (7.12). *Provided* $C_1(\ell(I)) < 1$, *the estimate*

$$\left|o_{I_1}^c(x) - o_{I_2}^c(x)\right|_E \leq |g_I(x)|_E + (1 - C_1(\ell(I)))^{-1} K_2 \overline{g}_I \tag{7.18}$$

holds.

Proof. Take the E-norm at both sides of (7.15) and note that this yields (7.6) with ϕ, h, and K replaced by, respectively, $\left|o_{I_1}^c(x) - o_{I_2}^c(x)\right|_E$, $|g_I(x)|_E$, and $\Lambda_{I_2}(x, \Omega_b)$. As (7.18) is nothing but (7.8) written out for this choice of ϕ, h and K, we are done (Strictly speaking we cannot apply Lemma 7.2 as formulated, since in that lemma the I-argument is the same for all functions, whereas now it differs. However, it should be clear that exactly the same sequence of arguments can be applied to yield (7.18)). ◻

It remains to combine the lemmas into a more informative statement.

Proposition 7.6 *Assume that there are positive constants* K_1 *and* K_2 *such that for all* $x \in \Omega$ *and all* $I \in \mathcal{B}(Z)$

$$|o_I(x)|_E \leq K_1, \tag{7.19}$$

$$\Lambda_I(x, \Omega_b) \leq K_2. \tag{7.20}$$

Assume, furthermore, that there are nondecreasing (as functions of s*) functions* $C_1(s)$, $C_2(x, s)$ *and* $C_3(x, s)$ *which tend to zero as* $s \downarrow 0$, *in the case of index* 2 *and* 3 *uniformly for* $x \in \Omega_b$ *(but not necessarily uniformly for* $x \in \Omega$*) such that*

$$\Lambda_I(x, \Omega_b) \leq C_1(\ell(I)), \tag{7.21}$$

$$|o_I(x) - o_J(x)|_E \leq C_2(x, \ell)\|I - J\| \tag{7.22}$$

$$V\left(\left|\Lambda_{\rho(\cdot)I}-\Lambda_{\rho(\cdot)J}\right|(x,\Omega_b)\right)\leq C_3(x,\ell)\|I-J\| \tag{7.23}$$

for all $x\in\Omega$ and all $I,\ J\in\mathcal{B}(Z)$ of equal length $\ell=\ell(I)=\ell(J)$. Then

$$|o_I^c(x)-o_J^c(x)|_E\leq\left(C_2(\ell)+C_4(x,\ell)\right)\|I-J\|, \tag{7.24}$$

where $C_2(s)$ and $C_4(x,s)$ are nondecreasing in s and tend to zero for $s\downarrow 0$ (in the case of C_4 pointwise for $x\in\Omega$ but uniformly for $x\in\Omega_b$).

Proof. By Lemma 7.3 and Lemma 7.4 we have

$$|g_I(x)|_E\leq C_4(x,\ell)\|I-J\|, \tag{7.25}$$

where

$$C_4(x,s):=C_2(x,s)+K_1\left(1+\left(1-C_1(s)\right)^{-1}K_2\right)C_3(x,s).$$

If we insert (7.25) into (7.18) we obtain (7.24) with

$$C_2(s)=\left(1-C_1(s)\right)^{-1}K_2\sup_{\xi\in\Omega_b}C_4(\xi,s).$$

□

As a straightforward corollary we obtain the following theorem.

Theorem 7.7 *Let $B_s(Z)$ be equipped with the supremum norm. Under the assumptions of Proposition* 7.6 *the pure mass action problem has a unique solution.*

Proof. It follows from the proposition that the inequality (7.1) holds, and this, as we have already shown, implies that the pure mass action problem has a unique solution. □

The derivation of the appropriate estimate for the L_1-norm proceeds along exactly the same lines. We start with the analogue of Lemma 7.3.

Lemma 7.8 *Assume that there are positive constants K_1 and K_2 and a nondecreasing function $C_1(s)$, with $\lim_{s\downarrow 0}C_1(s)=0$, such that for all $x\in\Omega$ and all $I\in\mathcal{B}(Z)$*

$$\int_{[0,\ell(I))}\left|o_{\rho(t)I}(x)\right|_E\,dt\leq K_1, \tag{7.26}$$

$$\Lambda_I(x,\Omega_b)\leq K_2 \tag{7.27}$$

and for all $x\in\Omega_b$ and all $I\in\mathcal{B}(Z)$

$$\Lambda_I(x,\Omega_b)\leq C_1(\ell(I)). \tag{7.28}$$

Then

$$\int_{[0,\ell(I))}\left|o_{\rho(t)I}^c\right|_E\,dt\leq K_1\left(1+\left(1-C_1(\ell(I))\right)^{-1}K_2\right) \tag{7.29}$$

for all $x\in\Omega$ and all $I\in\mathcal{B}(Z)$ with $C_1(\ell(I))<1$.

Proof. Consider again the convolution equation (6.5), but now in the form

$$o^c_{\rho(t)I} = o_{\rho(t)I} + (o^c * \Lambda)_{\rho(t)I} \,.$$

Taking the E-norm and integrating with respect to t over $[0, \ell(I))$ we obtain the inequality (7.6) with

$$\phi(I,x) = \int_{[0,\ell(I))} \left| o^c_{\rho(t)I}(x) \right|_E \, dt,$$

$$h(I,x) = \int_{[0,\ell(I))} \left| o_{\rho(t)I}(x) \right|_E \, dt,$$

$$K(I,x) = \Lambda_I(x, \Omega_b).$$

To see this, interchange the order of the two integrations in the convolution term. The inequality (7.29) is then obtained from (7.8) by using (7.26) – (7.28). ◻

In completely the same manner we can prove the analogue of Lemma 7.5.

Lemma 7.9 *Assume* (7.27) *and* (7.28). *Then for all* $I,\ J \in \mathcal{B}(Z)$ *of equal length* $\ell = \ell(I) = \ell(J)$ *with* $C_1(\ell) < 1$ *one has*

$$\int_{[0,\ell)} \left| o^c_{\rho(t)I}(x) - o^c_{\rho(t)J}(x) \right|_E \, dt \le \int_{[0,\ell(I))} \left| g_{\rho(t)I}(x) \right|_E \, dt +$$

$$(1 - C_1(\ell(I)))^{-1} \, K_2 \sup_{\xi \in \Omega_b,\, \alpha \in [0,\ell)} \int_{[0,\ell-\alpha)} \left| g_{\rho(t)\theta(-\alpha)I}(x) \right|_E \, dt. \tag{7.30}$$

Combining the lemmas 7.8 and 7.9 with Lemma 7.4 we obtain the following proposition and its more fundamental corollary.

Proposition 7.10 *Assume* (7.26) – (7.28) *as well as*

$$\int_{[0,\ell)} \left| o_{\rho(t)I}(x) - o_{\rho(t)J}(x) \right|_E \, dt \le C_2(x, \ell(I)) \|I - J\| \tag{7.31}$$

and (7.23) *for all* $I,\ J \in \mathcal{B}(Z)$ *of equal length* ℓ. *Then*

$$\int_{[0,\ell)} \left| o^c_{\rho(t)I}(x) - o^c_{\rho(t)J}(x) \right|_E \, dt \le \left(C_2(\ell) + C_4(x,\ell) \right) \|I - J\|, \tag{7.32}$$

where $C_2(s)$ *and, for every* $x \in \Omega$, $C_4(x,s)$ *are nondecreasing functions of* s *that tend to zero as* $s \downarrow 0$.

Theorem 7.11 *Let* $B_s(Z)$ *be equipped with the* L_1*-norm. Under the assumptions of Proposition* 7.10 *the pure mass action problem has a unique solution.*

In conclusion of this section we shall present the arguments that prepare the way for an application of the contraction mapping principle in the case of the generalized mass action problem.

Lemma 7.12 *Let* A_1 *and* A_2 *be Banach spaces and let for all* $a = (a_1, a_2),\ b = (b_1, b_2) \in A_1 \times A_2$, $H = (H_1, H_2) : A_1 \times A_2 \to A_1 \times A_2$ *satisfy*

$$\begin{aligned} \|H_1(a) - H_1(b)\|_{A_1} &\le \lambda \|a - b\|, \\ \|H_2(a) - H_2(b)\|_{A_2} &\le \lambda \|a - b\| + K \|a_1 - b_1\|_{A_1} \end{aligned}$$

for some positive constants λ *and* K. *Here* $\|\cdot\|$ *is the* l_1*-norm on* $A_1 \times A_2$. *Then*

$$\|H^2(a) - H^2(b)\| \le \left(4\lambda^2 + 3\lambda K \right) \|a - b\| \tag{7.33}$$

for all $a,\ b \in A_1 \times A_2$, *where* $H^2 = H \circ H$.

Proof. One has

$$\begin{aligned}\|H_1^2(a)-H_1^2(b)\| &\leq \lambda\|H(a)-H(b)\| \\ &\leq \lambda\left(2\lambda\|a-b\|+K\|a_1-b_1\|_{A_1}\right) \\ &\leq \Big(2\lambda^2+K\lambda\Big)\|a-b\|\end{aligned} \tag{7.34}$$

and

$$\begin{aligned}\|H_2^2(a)-H_2^2(b)\| &\leq \lambda\|H(a)-H(b)\|+K\|H_1(a)-H_1(b)\|_{A_1} \\ &\leq \Big(2\lambda^2+K\lambda\Big)\|a-b\|+K\lambda\|a-b\| \\ &= \Big(2\lambda^2+2K\lambda\Big)\|a-b\|.\end{aligned} \tag{7.35}$$

Adding (7.34) and (7.35) one obtains (7.33). □

Within our framework $A_i = B_s(Z_i)$ and H is the output map, cf. (6.3). The λ then depends on s and tends to zero as $s \downarrow 0$, while K stays bounded away from zero and infinity (so may be chosen independent of s). For s sufficiently small, $4\lambda^2 + 3\lambda K < 1$ and we can apply the contraction mapping theorem to H^2. The conclusion is that H^2 has a unique fixed point, say $\overline{a}$. But as $H(\overline{a})$ is a fixed point of H^2, too, uniqueness implies that actually $\overline{a}$ must be a fixed point of H itself.

In Lemma 7.12 we have chosen the l_1-norm on the product space $A_1 \times A_2$, but, as all norms on $\mathbf{R}^2$ are equivalent, any other choice would have done equally well. Of course the expression for the Lipschitz constant for H^2 in terms of λ and K depends on the choice, but in all cases this Lipschitz constant tends to zero as $\lambda \downarrow 0$.

8 Estimating individual output

Let us assume that the interaction variable takes values in a finite dimensional space, say $\mathbf{R}^k$. Moreover, let us specialise to the situation where the i-state space Ω is a (connected) subset of $\mathbf{R}^n$ for some n with piecewise smooth boundary. We now concentrate on *deterministic* development of individuals, which we call *growth*. We refer to (Diekmann *et al.* 1998, Section 8.3) for an example involving random movement in Ω.

Let $X_I(x_0)$ denote the i-state of an individual at time $\ell(I)$, given that

- it had i-state x_0 at time zero,
- it experienced input I,
- it survived.

Similarly, let $\mathcal{F}_I(x_0)$ denote the survival probability at time $\ell(I)$ of an individual which had i-state x_0 at time zero and experienced input I.

Concerning reproduction, let us assume that the state-at-birth has a distribution described by a probability measure m_b (concentrated on a subset Ω_b of Ω), irrespectively of the state of the mother at the moment of giving birth. The particular case of a fixed state-at-birth x_b corresponds to the choice $m_b = \delta_{x_b}$. Let $L_I(x_0)$denote the expected number of offspring produced by an individual with i-state x_0 at time zero in the time interval $[0, \ell(I))$ while experiencing input I.

The assumptions made above mean that the ingredients u_I and Λ_I take the forms

$$u_I(x,\omega) = \delta_{X_I(x)}(\omega)\mathcal{F}_I(x), \tag{8.1}$$

$$\Lambda_I(x,\omega) = L_I(x) m_b(\omega), \tag{8.2}$$

for $x \in \Omega$ and ω a measurable subset of Ω. As a consequence

$$o_I(x) = (\gamma \times u_I)(x) = \int_\Omega \gamma(\xi)\mathcal{F}_I(x)\delta_{X_I(x)}(d\xi) = \gamma(X_I(x))\mathcal{F}_I(x), \tag{8.3}$$

where γ is the individual output function.

We shall need the following hypotheses.

Hypothesis 8.1 There exists a constant K_2 and a nondecreasing function $C_1 : \mathbf{R}_+ \to \mathbf{R}_+$ tending to 0 as $s \downarrow 0$ such that for all $I \in \mathcal{B}(Z)$ and all $x \in \Omega$ one has

$$L_I(x) \leq K_2 \tag{8.4}$$

and for all $I \in \mathcal{B}(Z)$ and all $x \in \Omega_b$ one has

$$L_I(x) \leq C_1(\ell(I)). \tag{8.5}$$

Hypothesis 8.2 Let I and J be two inputs of equal lengths: $\ell(I) = \ell(J) =: \ell$. There exist finite positive numbers $C_X(\ell),\ C_{\mathcal{F}}(\ell), C_L(\ell)$, depending only on ℓ, such that for each $x_0 \in \Omega$

$$|X_I(x_0) - X_J(x_0)| \leq C_X(\ell)\int_0^\ell |I(s) - J(s)|ds, \tag{8.6}$$

$$|\mathcal{F}_I(x_0) - \mathcal{F}_J(x_0)| \leq C_{\mathcal{F}}(\ell)\int_0^\ell |I(s) - J(s)|ds, \tag{8.7}$$

$$|L_I(x_0) - L_J(x_0)| \leq C_L(\ell)\int_0^\ell |I(s) - J(s)|ds. \tag{8.8}$$

Hypothesis 8.1 expresses the natural requirement that no-one begets an infinite number of children and that newborns cannot get a positive number of offspring immediately upon birth. This latter requirement is of course automatically satisfied if every $x \in \Omega_b$ is a state with positive reproduction delay (cf. Definition 5.2).

Hypothesis 8.2 contains natural Lipschitz-type conditions, which, as we show below, can easily be verified if individual behaviour is described in terms of rates satisfying corresponding Lipschitz estimates.

Theorem 8.3 *Let $B_s(Z)$ be equipped with the supremum norm and let $\gamma : \Omega \to Z$ be bounded and globally Lipschitz continuous. Then, under Hypotheses* 8.1 *and* 8.2, *the pure mass action problem has a unique solution.*

Proof. According to Proposition 7.7 we have to verify that the inequalities (7.19) –(7.23) hold true. The estimate (7.19) holds because by (8.3) one has

$$|o_I(x)|_E = |\gamma(X_I(x))\mathcal{F}_I(x)|_E \leq \|\gamma\|_\infty .$$

It follows from (8.2), the fact that m_b is a probability measure, and Hypothesis 8.1 that (7.20) and (7.21) hold. One has

$$\begin{aligned} |o_I(x) - o_J(x)|_E &= |\gamma(X_I(x))\mathcal{F}_I(x) - \gamma(X_J(x))\mathcal{F}_J(x)|_E \\ &\leq K\,|X_I(x) - X_J(x)|_E + \|\gamma\|_\infty\,|\mathcal{F}_I(x) - \mathcal{F}_J(x)| \\ &\leq C(\ell)\int_{[0,\ell)} |I(s) - J(s)|ds \\ &\leq C(\ell)\ell\|I - J\|_\infty, \end{aligned}$$

which shows that (7.22) holds.

To prove (7.23), recall that for a function ϕ defined on an interval $[a, b]$ the total variation $V(\phi)$ is defined as

$$V(\phi) = \sup \sum_{j=1}^{n} [\phi(t_j) - \phi(t_{j-1})| ,$$

where the supremum is taken over all partitions $\{a = t_0, t_1, \ldots, t_{n-1}, t_n = b\}$ of $[a, b]$. One gets

$$\begin{aligned} & V\left(\left|L_{\rho(\cdot)I}(x) - L_{\rho(\cdot)J}(x)\right|\right) \\ = \ & \sup \sum \left|L_{\rho(t_j)I} - L_{\rho(t_j)J} - L_{\rho(t_{j-1})I} + L_{\rho(t_{j-1})J}\right| \\ = \ & \sup \sum \left|L_{\rho(t_j - t_{j-1})\theta(-t_{j-1})I} - L_{\rho(t_j - t_{j-1})\theta(-t_{j-1})J}\right| \\ \leq \ & \sup \sum C_L(t_j - t_{j-1}) \int_{[0, t_j - t_{j-1})} |\theta(-t_{j-1})I(s) - \theta(-t_{j-1})J(s)| \, ds \\ = \ & \sup \sum C_L(t_j - t_{j-1}) \int_{[t_{j-1}, t_j)} |I(s) - J(s)| ds \\ \leq \ & C_L(\ell) \int_{[0,\ell)} |I(s) - J(s)| ds, \end{aligned}$$

from which it follows that (7.23) is satisfied. □

The L^1-case is proven in a completely analogous manner by verifying that the assumptions of Proposition 7.10 hold. The assumption of a uniformly bounded life-expectancy is needed to verify (7.26). Therefore we formulate the following hypothesis.

Hypothesis 8.4 There exists an $M < \infty$ such that

$$\int_{[0,\ell(I))} u_{\rho(\sigma)I}(x, \Omega) d\sigma \leq M \tag{8.9}$$

for every $x \in \Omega$ and every $I \in \mathcal{B}(Z)$.

We state the result in the L^1-case without proof.

Theorem 8.5 *Let $B_s(Z)$ be equipped with the L^1-norm and let $\gamma : \Omega \to Z$ be bounded and globally Lipschitz continuous. Then, under Hypotheses* 8.1 — 8.4, *the pure mass action problem has a unique solution.*

Hypothesis 8.2 is easily verified if growth, survival and reproduction are modelled by instantaneous rates depending on the i-state and the environmental condition and if these rates are, for instance, globally Lipschitz continuous in both their variables. So let $g : \Omega \times Z \to \mathbf{R}_+$, $\mu : \Omega \times Z \to \mathbf{R}_+$, $\beta : \Omega \times Z \to \mathbf{R}_+$ be the growth, death and fecundity rate, respectively. This means that $t \mapsto X_{\rho(t)I}(x_0)$ is the unique solution of the initial value problem

$$\frac{d}{dt} x(t) = g(x(t), I(t)), \tag{8.10}$$

$$x(0) = x_0, \tag{8.11}$$

that

$$\mathcal{F}_I(x_0) = e^{-\int_0^{\ell(I)} \mu\left(X_{\rho(s)I}, I(s)\right) ds}, \tag{8.12}$$

and that

$$L_I(x_0) = \int_0^{\ell(I)} \beta\left(X_{\rho(s)I}(x_0), I(s)\right) \mathcal{F}_{\rho(s)I}(x_0) \, ds. \tag{8.13}$$

Proposition 8.6 *Assume that* $X_I(x_0),\ \mathcal{F}_I(x_0)$ *and* $L_I(x_0)$ *are defined by* (8.10) – (8.13), *where the functions* $g,\ \mu$ *and* β *are globally Lipschitz continuous in both variables and* β *is bounded. Then the Hypotheses* 8.1 *and* 8.2 *hold true. If* $\mu(x,z) \geq \varepsilon > 0$ *for all* $(x,z) \in \Omega \times Z$*, then Hypothesis* 8.4 *holds true.*

Proof. That Hypothesis 8.1 is satisfied follows under the given assumptions immediately from (8.13). By (8.1) and (8.12) one has

$$\int_{[0,\ell(I))} u_{\rho(\sigma)I}(x,\Omega)d\sigma = \int_{[0,\ell(I))} e^{-\int_0^\sigma \mu\left(X_{\rho(s)I}(x),I(s)\right)ds} d\sigma$$

from which Hypothesis 8.4 follows via the assumption made on μ.

To verify Hypothesis 8.2, first note that by (8.10) and (8.11) one has

$$X_I(x_0) = x_0 + \int_0^{\ell(I)} g\left(X_{\rho(s)I}, I(s)\right) ds$$

and hence, by the global Lipschitz continuity of g,

$$\left|X_{\rho(t)I}(x_0) - X_{\rho(t)J}(x_0)\right| \leq$$

$$K\int_0^t |I(s) - J(s)|\, ds + K\int_0^t \left|X_{\rho(s)I}(x_0) - X_{\rho(s)J}(x_0)\right| ds \tag{8.14}$$

for some finite constant K. Applying Gronwall's lemma to (8.14) one obtains

$$\mid X_I(x_0) - X_J(x_0) \mid \leq K\int_0^{\ell(I)} e^{K(\ell(I)-s)} |I(s) - J(s)|\, ds$$

from which (8.6) follows immediately.

Because $|e^{-x} - e^{-y}| \leq |x-y|$ for $x \geq 0,\ y \geq 0$, it follows from (8.12) and the global Lipschitz continuity of μ that

$$|\mathcal{F}_I(x_0) - \mathcal{F}_J(x_0)| \leq \int_0^{\ell(I)} \left|\mu\left(X_{\rho(s)I}, I(s)\right) - \mu\left(X_{\rho(s)J}, J(s)\right)\right| ds$$

$$\leq K\int_0^{\ell(I)} \left|X_{\rho(s)I}(x_0) - X_{\rho(s)J}(x_0)\right| ds + K\int_0^{\ell(I)} |I(s) - J(s)|\, ds. \tag{8.15}$$

(8.7) now follows from (8.6) and (8.15).

Finally, using the fact that $\mathcal{F}_I(x_0) \leq 1$ for all I, and the assumptions about β, one finds from (8.13) that

$$|L_I(x_0) - L_J(x_0)| \leq$$

$$\int_0^{\ell(I)} \left|\beta\left(X_{\rho(s)I}, I(s)\right) - \beta\left(X_{\rho(s)J}, J(s)\right)\right| \left|\mathcal{F}_{\rho(s)I}(x_0)\right| ds$$

$$+ \int_0^{\ell(I)} \left|\beta\left(X_{\rho(s)J}, J(s)\right)\right| \left|\mathcal{F}_{\rho(s)I}(x_0) - \mathcal{F}_{\rho(s)J}(x_0)\right| ds \leq$$

$$K\int_0^{\ell(I)} \left|X_{\rho(s)I}(x_0) - X_{\rho(s)J}(x_0)\right| ds + K\int_0^{\ell(I)} |I(s) - J(s)|\, ds$$

$$+K\int_0^{\ell(I)} \left|\mathcal{F}_{\rho(s)I}(x_0) - \mathcal{F}_{\rho(s)J}(x_0)\right| ds. \tag{8.16}$$

(8.8) follows from (8.6), (8.7), and (8.16). □

As we have argued in Section 1, certain idealisations, which are made to keep the model parameter scarce, yield functions γ that have jumps (and so are only piecewise Lipschitz continuous). The aim of the remaining part of this section is to derive the estimate (7.31) for a simple prototype example of a γ with jumps.

Let us assume that the i-state space is one-dimensional, that is, $\Omega \subset \mathbf{R}$. We ignore the possibility of death and assume that the individual growth rate g is bounded away from zero, that is, there exists an $\varepsilon > 0$, such that $g(x,z) \geq \varepsilon$ for all $(x,z) \in \Omega \times Z$ (see Remark 8.8 if you find this assumption overly restrictive). Finally, let there be a jump point $\overline{x} \in \Omega$ such that

$$\gamma(x) = \begin{cases} 0 & \text{if } x < \overline{x}, \\ 1 & \text{if } \overline{x} < x, \end{cases}$$

the value of γ at $\overline{x}$ being irrelevant.

For a given input I and given initial i-state x_0 we can ask when an individual will reach the jump point $\overline{x}$. The answer is obtained by solving the equation

$$X_{\rho(s)I}(x_0) = \overline{x} \tag{8.17}$$

for s as a function of x_0 and I. There may be no solution, but if there is one, it is unique by the strict monotonicity of the map $s \mapsto X_{\rho(s)I}(x_0)$. We denote the solution (defined on a subset of $\Omega \times \mathcal{B}(Z)$ and taking values in $\mathbf{R}_+$) by $\tilde{s} = \tilde{s}(x_0, I)$.

Alternatively we may solve (8.17) for x_0 as a function of s and I. The solution (defined on a subset of $\mathbf{R}_+ \times \mathcal{B}(Z)$ and taking values in Ω) is denoted by $\tilde{x} = \tilde{x}(s, I)$.

Now let I and J be inputs of equal length $\ell(I) = \ell(J) = \ell$. Then we define

$$\begin{aligned} \tilde{s}_{\min}(x_0) &= \min\{\tilde{s}(x_0, I), \tilde{s}(x_0, J)\}, \\ \tilde{s}_{\max}(x_0) &= \max\{\tilde{s}(x_0, I), \tilde{s}(x_0, J)\} \end{aligned}$$

with the conventions that $\tilde{s}_{\max}(x_0) = \ell$ if at least one of the elements (x_0, I), (x_0, J) is not in the domain of $\tilde{s}$ and that $\tilde{s}_{\min}(x_0) = \ell$ if both these elements are not in the domain of $\tilde{s}$. We need these quantities to describe the function

$$\psi(x_0, t) = \left|\gamma\left(X_{\rho(t)I}(x_0)\right) - \gamma\left(X_{\rho(t)J}(x_0)\right)\right| \tag{8.18}$$

which is at the centre of our interest because

$$\left|o_{\rho(t)I}(x_0) - o_{\rho(t)J}(x_0)\right| = \psi(x_0, t). \tag{8.19}$$

Clearly,

$$\psi(x_0, t) = \begin{cases} 0 & \text{if } 0 \leq t < \tilde{s}(x_0), \\ 1 & \text{if } \tilde{s}_{\min}(x_0) < t < \tilde{s}_{\max}(x_0), \\ 0 & \text{if } \tilde{s}_{\max}(x_0) < t \end{cases} \tag{8.20}$$

and consequently

$$\int_o^\ell \psi(x_0, t)\, dt \leq \tilde{s}_{\max}(x_0) - \tilde{s}_{\min}(x_0). \tag{8.21}$$

We now claim that

$$\tilde{s}_{\max}(x_0) - \tilde{s}_{\min}(x_0) \leq \frac{C(\ell))}{\varepsilon} \int_0^\ell |I(t) - J(t)|\, dt. \tag{8.22}$$

To substantiate this claim we assume that $\tilde{s}_{\min}(x_0) = \tilde{s}(x_0, I)$. The lower bound of the growth rate implies that for $t \geq \tilde{s}(x_0, I)$ we have

$$X_{\rho(t)I}(x_0) \geq \overline{x} + \varepsilon(t - \tilde{s}(x_0, I)).$$

On the other hand, we have the Lipschitz estimate (8.6) which implies that

$$\left|X_{\rho(t)I}(x_0) - X_{\rho(t)J}(x_0)\right| \le C \int_0^\ell |I(s) - J(s)|\, ds.$$

With $\tilde{s} = \tilde{s}_{\max}(x_0) = \tilde{s}(x_0, J)$ for brevity, we have

$$\begin{aligned}\overline{x} = X_{\rho(\tilde{s})J}(x_0) &= X_{\rho(\tilde{s})I}(x_0) + X_{\rho(\tilde{s})J}(x_0) - X_{\rho(\tilde{s})I}(x_0)\\ &\ge \overline{x} + \varepsilon(\tilde{s} - \tilde{s}(x_0, I)) - C \int_0^\ell |I(s) - J(s)|\, ds,\end{aligned}$$

which implies that

$$\tilde{s}(x_0, J) - \tilde{s}(x_0, I) \le \frac{C}{\varepsilon} \int_0^\ell |I(s) - J(s)|\, ds,$$

that is, the estimate (8.22) holds.

In the estimate (8.22), however, we lose a lot of information. Indeed, $\tilde{s}_{\max}(x_0) - \tilde{s}_{\min}(x_0) = 0$ when both $\tilde{s}_{\max}(x_0)$ and $\tilde{s}_{\min}(x_0)$ are equal to ℓ, so in particular when $x_0 \le \min\{\tilde{x}(\ell, I), \tilde{x}(\ell, J)\}$. If we combine this observation with the estimate (8.22) we can deduce from (8.21) the estimate

$$\int_0^\ell \psi(x_0, t)\, dt \le \frac{C(\ell)}{\varepsilon} \int_0^\ell |I(s) - J(s)|\, ds \chi_{[0,\infty)}(x_0 - \min\{\tilde{x}(\ell, I), \tilde{x}(\ell, J)\}), \tag{8.23}$$

where $\chi_{[0,\infty)}$ is the characteristic function of $[0, \infty)$, that is, the Heaviside function. Recalling (8.19) we note that this is exactly of the form (7.31) with $C_2(x_0, \ell)$ being, for fixed $x_0 < \overline{x}$, equal to zero for ℓ sufficiently (depending on x_0) small.

Proposition 8.7 *Let $\Omega \subset \mathbf{R}$ and $Z \subset \mathbf{R}^k$ and let $\gamma : \Omega \to Z$ be piecewise globally Lipschitz continuous. Assume that $\varepsilon > 0$ exists such that for all $I \in \mathcal{B}(Z)$ and all $x_0 \in \Omega$ the inequality*

$$X_I(x_0) - x_0 \ge \varepsilon \ell(I)$$

holds. Moreover, let the Lipschitz estimates (8.6) *and* (8.7) *hold. Then, if I and J are two inputs of equal length ℓ, we have the inequality*

$$\int_0^\ell \left|\gamma\left(X_{\rho(t)I}(x_0)\right) \mathcal{F}_{\rho(t)I}(x_0) - \gamma\left(X_{\rho(t)J}(x_0)\right) \mathcal{F}_{\rho(t)J}(x_0)\right| dt \le$$

$$C(x_0, \ell) \int_0^\ell |I(t) - J(t)|\, dt$$

for a function C for which $\lim_{\ell \downarrow 0} C(x_0, \ell) = 0$ for every $x_0 \in \Omega$.

This proposition can easily be proven by using the estimate (8.23) and the fact that a piecewise Lipschitz continuous function can be written as the sum of a truly Lipschitz continuous function and a finite number of multiples of Heaviside functions.

Remark 8.8 Note that in a similar manner one can relax the lower bound on the growth rate: it need only hold near to the jump points of γ.

To conclude, we stress the two points that are essential for dealing successfully with discontinuous functions γ (‘successfully’ meaning that we can use a contraction mapping argument to prove well-posedness). The first is that we use the L_1-norm to measure inputs and outputs (indeed, (8.20) shows that the supremum norm of the difference in output cannot be bounded by a multiple of the sup-norm of the difference in input). The second is that it is sufficient that the Lipschitz constants tend to zero pointwise, but not necessarily uniformly, in the i-state x_0, when the time window shrinks to zero.

9 Back to the cannibalism example

In this section we show how the general theory applies to a nontrivial example, viz. a model involving cannibalistic behaviour. In Section 1 we introduced various ingredients of such a model, but we did not provide a full specification. So before embarking upon the application of our results, we first give a more precise description.

Individuals are characterized by their size $x \geq x_b$, where x_b is the size at which they are born. They grow, die and reproduce with rates g, μ and β, respectively. These rates depend on the i-state as well as on the environmental condition.

The pde formulation of the model is

$$\begin{aligned} \frac{\partial n}{\partial t} + \frac{\partial}{\partial x}(gn) &= -\mu n, \\ gn\big|_{x=x_b} &= \int_{x \geq x_b} \beta n dx, \end{aligned}$$

with g, μ, and β as specified below. In our view, this is only a convenient short-hand notation. In a preprocessing step we form u_I and Λ_I via (8.1) and (8.2) with X_I, $\mathcal{F}_I$, and L_I given by (8.10) — (8.13). Next we apply the machinery developed in Part I (Diekmann *et al.* 1998) and the present paper. The main result is that a population semiflow is constructively defined, given certain assumptions on g, μ, and β. In our elaboration below we do not strive for the utmost generality. Yet, on the other hand, we want to demonstrate the flexibility of our approach by including a case in which the behaviour of individuals changes abruptly upon passing a critical size.

The environmental condition has three components corresponding to, respectively, the reduction factor $(I_3(t,x))^{-1}$ of search time due to handling of prey, the rate of food ingestion $I_1(t,x)$ expressed in energy units, and the death rate $I_2(t,x)$ partly due to cannibalism. With slight abuse of notation we now formulate an assumption concerning the specific form of these quantities as follows:

$$I_3(t,x) = 1 + H(x)C(x)Z(x) + h_1(x)\psi(x)\tilde{O}_3(t), \tag{9.1}$$

$$I_1(t,x) = \frac{E(x)C(x)Z(x) + \psi(x)\tilde{O}_1(t)}{I_3(t,x)}, \tag{9.2}$$

$$I_2(t,x) = \sigma(x) + \phi(x)\tilde{O}_2(t). \tag{9.3}$$

The meaning of $H(x)$, $C(x)$, $Z(x)$, $\psi(x)$, $E(x)$, $\sigma(x)$ and $\phi(x)$ as well as that of $\tilde{O}_1(t)$ and $\tilde{O}_2(t)$ has already been explained in Section 1. The additional assumption underlying (9.1) – (9.3) is that

$$h(x,y) = h_1(x)h_2(y) \tag{9.4}$$

and that

$$\tilde{O}_3(t) = \int_{[x_b,\infty)} h_2(y)\phi(y)m_t(dy). \tag{9.5}$$

There are two more parameters entering the model description. One is the maintenance rate constant ζ. The second is the size specific allocation rule $\kappa(x)$ which describes how much of the ingested energy goes to growth and how much to reproduction.

We are now ready to give the formulas for g, μ, β and the three-vector γ in terms of x and the three-vector $I(t)$, that together fully specify the model:

$$g(x,I(t)) = (1-\kappa(x))\frac{E(x)C(x)Z(x) + \psi(x)I_1(t)}{1 + H(x)C(x)Z(x) + h_1(x)\psi(x)I_3(t)} - \zeta x, \tag{9.6}$$

$$\mu(x,I(t)) = \sigma(x) + \phi(x)I_2(t), \tag{9.7}$$

$$\beta(x, I(t)) = \kappa(x) \frac{E(x)C(x)Z(x) + \psi(x)I_1(t)}{1 + H(x)C(x)Z(x) + h_1(x)\psi(x)I_3(t)}, \tag{9.8}$$

$$\gamma_1(x) = e(x)\phi(x), \tag{9.9}$$

$$\gamma_2(x, I(t)) = \frac{\psi(x)}{1 + H(x)C(x)Z(x) + h_1(x)\psi(x)I_3(t)}, \tag{9.10}$$

$$\gamma_3(x) = h_2(x)\phi(x) \tag{9.11}$$

All functions featuring in this description take nonnegative values. We assume that for some $\varepsilon > 0$ and for all $x \geq x_b$

$$(1 - \kappa(x)) \frac{E(x)C(x)Z(x)}{1 + H(x)C(x)Z(x)} - \zeta x \geq \varepsilon, \tag{9.12}$$

which tells us that growth will never stop (in fact, this assumption is debatable and alternatives like von Bertalanfy growth and/or a reserve compartment have been considered, cf. (Kooijman 2000; Metz and Diekmann 1986); however, here we do not want to complicate the formulation of the results by having a size upper bound and the possibility of shrinking when maintenance cannot be covered by food). We also assume that all functions of x are bounded.

Theorem 9.1 *Let* $\zeta > 0$ *and let bounded, nonnegative functions* E, C, Z, e, ϕ, ψ, σ, H, h_1, h_2, *and* κ, *defined on* $[x_b, \infty)$, *be given. Assume that* (9.12) *holds. Also assume that all functions are globally Lipschitz continuous, with exception of* ψ, *which is only piecewise globally Lipschitz continuous. Then there exists a population semiflow corresponding to the individual behaviour as embodied in* (9.6) — (9.11).

Sketch of proof. When $h_1(x) = 0$ (that is, when cannibalistic predation has negligible influence on search time) and all functions of x are globally Lipschitz continuous, we can apply Theorem 8.3. Retaining the Lipschitz condition but allowing $h_1(x)$ to be nontrivial, we have to extend the underlying lemmas and theorems by means of Lemma 7.12. If we choose, for instance, $\psi(x) = \chi_{[0,\infty)}(x - x_A)$ we need Theorem 7.11 in combination with Proposition 8.7. □

10 Concluding remarks

In this paper we have proven existence and uniqueness of solutions of a general nonlinear structured population model and applied the result to a concrete model involving cannibalistic behaviour. We trust that our approach is such that it applies directly, or with only slight modifications, to a large class of structured population models.

A characteristic feature of structured population models is that the nonlinearity enters the model via feedback through the environment. This fact gives a clue to the existence and uniqueness proof: One first pretends that the environmental condition (the input) is known during a time-interval, then one calculates the corresponding output and iterates. The solution is thus constructed by successive approximations. In the context of structured population models this idea goes back (in the case of age-structured models) at least as far as Gurtin and MacCamy (1974). Various extensions and generalizations of the Gurtin-MacCamy model have been treated by essentially the same method in a number of papers; see the book by Webb (1985) and the references therein.

Age-dependent problems are very special in the sense that aging is not affected by the environment: chronological age always advances at the same rate as time. Nonlinear age-structured models are thus *semi-linear* problems, which are rather innocent nonlinear perturbations of a well-understood linear problem. When the individual development rate is allowed to depend on the environmental input, the problem becomes quasi-linear and thus essentially more difficult. Consequently there are only a few papers with existence and uniqueness proofs for such models, the most important being (Tucker and Zimmermann 1988; Thieme 1988; Calsina and Saldaña 1995, 1997). Tucker and Zimmermann (1988) assumed that the state-at-birth is distributed and that the population can be described by a density function; Thieme (1988) concentrated on the Kooijman-Metz *Daphnia* model and related certain model assumptions concerning individual energy allocation to uniqueness of solutions; Calsina and Saldaña (1995, 1997) did restrict to one-dimensional i-state space, in other words, to size structured populations.

All the authors mentioned above formulated their models analogously with the age-structured model as a hyperbolic partial differential equation supplemented by a nonlocal boundary condition describing the birth process. Diekmann *et al.* (2000) gave examples of how uniqueness can fail for such equations and pointed out that the problems leading to nonuniqueness are completely hidden in the pde formulation (see also Diekmann *et al.* 1993a, 1995). Therefore we have in this paper chosen the "cumulative" formulation of structured population models (Diekmann *et al.* 1993b, 1998), which takes as model ingredients not the individual vital rates, but the kernels u_I and Λ_I. An additional bonus of this approach is that stochasticity at the individual level can be incorporated at no extra cost.

Next on our agenda is the writing of a paper showing how to determine in an efficient manner steady p-states from the ingredients u_I, Λ_I and γ. This is essentially an elaboration of Theorem 6.1 in (Diekmann *et al.* 1998) together with a feedback fixed point problem. A formulation of a linearized stability test in terms of the position of the roots of a characteristic equation in the complex plane relative to the imaginary axis seems within reach (see e.g. Kirkilionis *et al.*, preprint). A rigorous justification of this test, however, is still a daunting task.

Finally we emphasize that our approach is not restricted to single-species models. As formulated in this paper our model actually includes the multi-species case: If there are k interacting species with individual state spaces, $\Omega_1, \Omega_2, \ldots, \Omega_k$, respectively, then one simply defines $\Omega = \Omega_1 \bigcup \Omega_2 \bigcup \ldots \bigcup \Omega_k$. The species interactions are modelled in terms of the environmental inputs.

References

Calsina, À. and Saldaña, J. (1995). A model of physiologically structured population dynamics with a nonlinear individual growth rate, *J. Math. Biol.* 33:335–364.

Calsina, À. and Saldaña, J. (1997). Asymptotic behaviour of a model of hierarchically structured population dynamics, *J. Math. Biol.* 35:967–987.

Diekmann, O., Gyllenberg, M., and Thieme, H.R. (1993a) Perturbing semigroups by solving Stieltjes renewal equations. *Differential and Integral Equations* 6: 155–181.

Diekmann,O., Gyllenberg, M., Metz, J.A.J. and Thieme, H.R. (1993). The "cumulative" formulation of (physiologically) structured population models, In *Evolution Equations, Control Theory and Biomathematics*, Ph. Clément and G. Lumer (Eds.), Marcel Dekker, New York, pp. 145 - 154.

Diekmann, O., Gyllenberg, M., and Thieme, H.R. (1995) Perturbing evolutionary systems by step responses and cumulative outputs. *Differential and Integral Equations* 8: 1205–1244.

Diekmann, O., Gyllenberg, M., Metz, J.A.J., and Thieme, H.R. (1998). On the formulation and analysis of general deterministic structured population models: I Linear theory. *J. Math. Biol.* 36: 349-388.

Diekmann, O., Gyllenberg, M., and Thieme, H.R. (2000). Lack of uniqueness in transport equations with a nonlocal nonlinearity, *Mathematical Models and Methods in Applied Sciences* 10: 581-592.

Dieudonné, J. (1969) *Foundations of Modern Analysis*, Enlarged and Corrected Printing, Academic Press, New York.

Gurtin, M.E. and MacCamy, R.C. (1974). Non-linear age-dependent population dynamics, *Arch. Rat. Mech. Anal.* 54: 281–300.

Kelley, J.L. and Namioka, I. (1963). *Linear Topological Spaces*, Van Nostrand, Princeton, N.J.

Kirkilionis, M., Diekmann, O., Lisser, B., Nool, M., de Roos, A.M., and Sommeijer, B. (preprint). Numerical continuation of equilibria of physiologically structured population models. I. Theory.

Kooijman, S.A.L.M. (2000). *Dynamic Energy and Mass Budgets in Biological Systems*, Cambridge University Press, Cambridge.

Metz, J.A.J. and Diekmann, O. (1986). *The Dynamics of Physiologically Structured Populations*. Lecture Notes in Biomathematics 68. Springer, Berlin.

H.R. Thieme (1988). Well-posedness of physiologically structured population models for *Daphnia magna*, *Journal of Mathematical Biology* 26: 299–317.

Tucker and Zimmermann (1988). A nonlinear model of population dynamics containing an arbitrary number of continuous structure variables, *SIAM J. Appl. Math.* 48: 549–591.

Webb, G.F. (1985) *Nonlinear Age-Dependent Population Dynamics*, Marcel Dekker, New York.