IV Age Dependence

J.A.J. Metz & O. Diekmann

1. Age as a substitute for comprehension

1.1. Why this special attention

On the whole the theory of structured population models is still *in statu nascendi.* We have a firm idea where the linear theory is heading, but a great deal of work remains to be done to get even a semblance of completeness, and our present understanding of nonlinear problems is scanty at best (but developing rapidly!). However, there is one specific area that is already well past puberty: that of purely age dependent problems. The deterministic linear theory, based on the so-called linear renewal (or Lotka's integral) equation, started with SHARPE's and LOTKA's work in 1911, still forms the backbone of human demography (KEYFITZ, 1968; COALE, 1972; POLLARD, 1973; a history of the early stages of the theory may be found in SAMUELSON, 1976), a corresponding stochastic theory of age dependent branching processes is nearing completion (JAGERS, 1975), and a great number of papers and one whole book (WEBB, 1985a) deal with nonlinear extensions.

It is certainly not our intention to add to this burgeoning specialist literature here especially since age is a variable with only minimal physiological connotations (see below and/or the discussion in III.2.2). However, as it turns out that often the easiest way to analyse a structured population model, at least in the linear case, is by transforming it into an equivalent renewal equation, we feel obliged to say at least a minimal amount about this technical device. In addition this gives us the opportunity to discuss in a concrete case how a special biological feature like "all individuals are born equal" can be exploited to arrive at clearcut results.

In this section we shall discuss the characterization of problems that allow an age type representation, either derived directly from observable i-behaviour or from some other more complicated model, as well as the related topic of using integral equations (as opposed to partial differential equations) as a modelling tool. The first section by necessity is fairly philosophical. The remaining sections will be more technical. In section 2 we shall introduce the basic linear theory centering around the so-called renewal theorem. In section 3 we shall consider some applications of that theorem in more complicated situations, and in section 4 we shall discuss some non-linear extensions. In these sections we shall be frankly parochial in that we concentrate mainly on examples from our own work. Section 5, finally, and also subsection 2.3.3 will be devoted to a totally different topic, the reduction of age dependent models to p-behaviourally equivalent differential equations in $\mathbb{R}^k$, $k<\infty$. Comparable techniques apply to other types of structured population models, but in the age dependent case we have by now a particularly good idea about their extent.

1.2. Which problems allow an age representation?

Two special biological properties set problems allowing an age representation apart from any other population problems. These are

(1) the population dynamical behaviour of any individual is in no way related directly to that of its parent(s) or the state of the environment immediately preceding its birth

(2) the average population dynamical properties of an individual, like mean rate of giving birth or probability of dying, do not depend on which circumstances it has experienced in its past. (NB There is no objection against a dependence of its birth and death rates on the current environment but the effect should be instantaneous.)

REMARK 1.2.1: We are referring here to what might be called generalized or long term population dynamically equivalent representations of the age type. This in contrast to strict i-state space representations which are supposed to

represent faithfully the actual *i*-behaviour in all its fine probabilistic detail. The difference is best explained by an example.
If a female human has just given birth we know that it will take her at least nine months before she can do so again. Therefore the information whether she is in the nine month *post partem* window or not allows us a better prediction of her future probabilities of giving birth. By the same argument the population dynamical future of a population artificially started with women who just gave birth should differ from that of a population with the same age distributions but in which none of the women are in the *post partem* window. However, under constant circumstances at least, the fraction of women in the *post partem* window in any cohort born into the population will just be a fixed function of the age of that cohort and so will be the *per capita* contribution of that cohort to the population birth rate. This *per capita* cohort birth rate should be calculated by taking the average over (the probability measure on the space of) all possible individual parturition histories.
An age representation will be strict if and only if in addition to conditions (1) and (2) we have that: (3) the individual birth process has so-called independent increments (which in the simplest case when multiple births are excluded means that it is a Poisson process) and the event of death is independent of the birth events, except for the possibility of a burst of births coinciding with the death event.

Properties (1) and (2) have counterparts referring to possible i-state representations of our individuals:

(1′) all individuals start identical in the sense that the probability distribution of their state at birth is identical, and

(2′) once born an individual's state process unfolds totally independent of its environment.

Here (2′) should be taken to imply that the differences between death rates in states that are attainable at the same age should remain constant at all time, for otherwise the environmental history may exert an influence on the relative likeliness of the various alternative *i*-state histories through selective deaths (see exercise 1.4.6 for an example). It can easily be seen that properties (1′) and (2′) imply conditions (1) and (2) respectively. Strictly speaking we cannot assert the reverse implication. However, as indicated in the next remark the counter examples are necessarily rather artificial.

REMARK 1.2.2: A sufficient condition for the reverse implication to hold true is that (a) the age representation is strict (see remark 1.2.1), and (b) we restrict the attention to minimal *i*-state representations (see III.2.2). Condition (a) in essence excludes the possibility that the environmental history experienced by an individual influences the fine detail of its actual birth process but not its average birth rate, and (b) in essence excludes the possibility that there exist components of the *i*-state vector which are for example related to the state of an individual's parents but which do not influence behaviour, as well as the presence of potential environmental influences on *i*-state behaviour in non-reachable parts of the *i*-state space.

If the physical environment is constant and there also is no direct interaction between individuals, condition (2) is satisfied by default. So for constant environmental conditions a great deal more problems allow an age representation than when conditions vary. (The *Daphnia* model from I.3 provides an example.)

REMARK 1.2.3: In the biological literature one frequently encounters a generalization of the age concept called physiological age. The most general and not very useful interpretation is that (i) an individual's average population dynamical properties can be represented fully by a one dimensional state process where the state variable can never decrease, and (ii) the state variable at birth can be set equal to zero (Note that when (i) is fulfilled (ii) is just equivalent to our earlier condition (1).) In this sense the *Daphnia* model from I.3 allows a physiological age representation.
In a more restricted interpretation the speed at which physiological age changes should be independent of physiological age itself. If we have some other *i*-state representation, say based on physiological insights, then for a physiological age representation to be possible it is sufficient that (1′) holds and that:

(2″) the only external influence allowed on the development of the *i*-state is through a multiplicative action on the speed of the *i*-state process as a whole.

(When the physiological age representation is strict and the other *i*-state representation is minimal (compare remark 1.2.2), these two conditions are also necessary.) If this multiplicative factor is called v, then we can introduce the function

$$p(t) := \int_0^t v(\tau)d\tau \tag{1.2.1}$$

called physiological time, and calculate the present physiological age a_p of an individual born at t_b as

$$a_p = p(t) - p(t_b) \tag{1.2.2}$$

(Compare the papers by Gurney, Nisbet & Blythe and by Goudriaan in part B.)

1.3. Integral equations as a natural modelling tool

For constant environmental conditions the observation that all individuals are born alike allows the derivation from first principles of an integral equation for the (total or population) birth rate b, alternatively called Lotka's equation by demographers and the renewal equation by mathematicians. Let $\phi(a)$ denote the mean number of offspring that an individual will beget per unit of time, a time units after it is born (dead individuals being assumed to have zero birth rate), and $g(t)$ the number of births per unit of time into the population which are not daughters of individuals born after $t=0$ *(i.e.* direct offspring from the founder population or possibly births from outside sources). Then the overall birth rate b can be written as the sum of g and the cumulation of the contributions from all individuals born after time zero:

$$b(t) = \int_0^t b(t-a)\phi(a)da + g(t)\,. \tag{1.3.1}$$

The functions ϕ and g are known as the birth kernel and the forcing function respectively. Equation (1.3.1) has been the object of detailed mathematical study for a long time. Section 2 will give a rough summary of the main results. Subsection 1.4 will give some examples of the calculation of the birth kernel ϕ and the forcing function g.

Given the birth function b it is possible to calculate any linear functional *(i.e.* weighted cumulation) of the population state, like total number of individuals alive, or total biomass, as a function of time. For example, if $\mathfrak{P}(a)$ denotes the probability that an individual survives at least to age a then the total population size N at time t equals

$$N(t) = \int_0^t b(t-a)\mathfrak{P}(a)da + M(t)\,, \tag{1.3.2}$$

where $M(t)$ denotes the number of survivors from the founder population.

If the environmental conditions vary dependent on the population development, then in the most general case there probably is no good alternative to writing down a full partial differential equation for the age distribution, along the lines laid out in chapters I and III. However, in many specific cases, for example when the total birth output is modified through mutual interaction of the neonates or by interaction with some specific segment of the population the size of which can be calculated as a linear functional of the birth history, it is still possible to derive a, now essentially non-linear, integral equation akin to equation (1.3.1), either directly or through the indirect route of first writing down a partial differential equation and then integrating it along the characteristics.

EXAMPLE 1.3.1: In III.6.2, example 6.2.4, we derived the following equation for the relation between the population birth rate b and recruitment h resulting from nursery competition

$$h = be^{-F(m)} \tag{1.3.3}$$

$$m = (b / F(m))(1-e^{-F(m)}) \tag{1.3.4}$$

where F is a function which tells how death rate in the nursery depends on crowding m, *i.e.* the number of competitors multiplied with their competitive strength. Therefore, if there are no other density dependent effects,

$$b(t) = \int_o^t \phi(a)h(b(t-a))da + g(t) \tag{1.3.5}$$

where $h(b)$ is defined by (1.3.3) with m the solution of (1.3.4). The dynamic properties of (1.3.3) - (1.3.5) will be examined in VI.3.

EXERCISE 1.3.2: Derive (1.3.5) by integrating the partial differential equation part of (III.6.2.9) along the characteristics.

The following example shows how equations like (1.3.1) and (1.3.5) can be derived direct from first principles in a slightly more rigorous manner. The main interest of this derivation is that the step of replacing detailed individual processes by just the average *pro capita* birth rate instead of being relegated to the verbal preparatory stages emerges as a consequence of the calculations.

EXAMPLE 1.3.3: *Kermack's and McKendrick's general epidemic* In 1927 KERMACK and MCKENDRICK introduced a general epidemic model in which infectivity of an ill individual was assumed to depend in some general way on the time elapsed since it was infected. Further assumptions of the model are that all susceptibles are equally vulnerable to infection, and that the number of susceptibles can only alter due to them being infected, *i.e.* there are no new susceptibles born to the population, susceptibles do not die (ill individuals may) and there is no return to susceptibility at any time after an individual has been infected. Apart from that only the usual law of mass action assumption is made. Below we shall derive their basic equation from the most general assumptions along the lines laid out in METZ (1978).

As the course of illness v of a particular individual is assumed not to depend on either the way it got infected or on the circumstances the individual experiences after its infection we may proceed as if each individual is just labeled by the course its illness will take if ever it gets infected. If we wish we may think of any variability between individual courses of illness as being caused by genuine differences between the individuals themselves (provided those differences are in no way related to an individual's proneness to contracting an infection!) Let V denote the set of possible courses of individual illnesses. Then we can conceive our population of infected individuals, alive, recovered or dead, as a frequency distribution over this set. As V will in general be a very complicated set there is no direct way to represent our population of infecteds as a density over V. Therefore we shall represent it as a measure, to be denoted as m, *i.e.* if U denotes some sufficiently well behaved subset of V then $m(t)$ attributes to U just the number $m(t,U)$ (or rather the number per unit of area, but we shall omit this qualification from now on) of individuals infected up to time t whose courses of illness (are destined to) lie in U. Let $a(v,\tau)$ denote the infectivity contributed by an individual whose course of illness is v and who was infected τ time units ago (we assume that our parametrization is sufficiently detailed that an individual's infectivity depends deterministically on v and τ). Then the total infectivity y at time t equals

$$y(t) = \int_0^t \int_V a(v,\tau)\dot{m}(t-\tau,\{dv\})d\tau + g(t) \tag{1.3.6}$$

where $\dot{m} = \frac{dm}{dt}$, $m(t,\{dv\})$ denotes the number of infecteds up to time t whose course of illness lie in an infinitesimal set $\{dv\}$ around the point v, and g denotes the infectivity due to individuals already ill at $t = 0$ or derived from outside sources. If s is the density of susceptibles then the law of mass action gives

$$\dot{s} = -ys\,. \tag{1.3.7}$$

Finally our assumption that there is no relation between an individual organism's susceptibility to infection and the course of its illness if ever it gets infected gives us

$$\dot{m}(U) = -\dot{s}\mu(U) \tag{1.3.8}$$

where μ is the probability that an as yet uninfected organism's course of illness is destined to lie in U. Substituting (1.3.8) in (1.3.6) we get

$$y(t) = -\int_0^t \dot{s}(t-\tau)\int_V \mu(\{dv\})a(v,\tau)d\tau + g(t) \tag{1.3.9}$$

or

$$y(t) = -\int_0^t \dot{s}(t-\tau)\phi(\tau)d\tau + g(t) \tag{1.3.10a}$$

with

$$\phi(\tau) = \int_V a(v,\tau)\mu(\{dv\})\,. \tag{1.3.11}$$

Recalling that an average is nothing but the integral of a random variable over the corresponding probability measure, we see that ϕ is just the average infectivity of an individual who has been infected τ time units ago. And recalling (1.3.7) we see that our epidemic process can be modelled by (1.3.10a) together with

$$\dot{s}(t) = -y(t)s = -y(t)\left(\int_0^t \dot{s}(t-\tau)d\tau + s_0\right). \tag{1.3.10b}$$

(1.3.10) is again a non-linear integral equation, this time relating the value of $\dot{s}$ at t to the behaviour of $\dot{s}$ during the interval $(0,t)$. In section 4 we shall deal in depth with the wealth of results that this equation permits one to derive. A related model in which infected individuals eventually return to the susceptibility class is treated in VI.2.

REMARK 1.3.4: Kermack's and McKendrick's paper is the classic of the mathematical theory of epidemics. Unfortunately it appears to be read considerably less frequently than it is cited. Apparently the idea got around that the paper only deals with an exceedingly special case which allows a simple ordinary differential equation representation in $\mathbb{R}^2$, of the kind discussed in 5.2. This is probably due to the following unfortunate combination of causes: a) the general equations proposed by Kermack & McKendrick were not in the mainstream of applied mathematics for some time to come, b) in his interesting and influential 1956 paper dealing with stochastic extensions of the simple model Kendall somehow gave the impression that this simple model essentially was what the Kermack & McKendrick paper was about, c) the original paper is not overly easy to obtain (this has recently been redressed by the reprint in OLIVEIRO PINTO & CONOLLY (1981)). As a result up to the present day papers get published purporting to analyse extensions of the Kermack & McKendrick model, whereas in fact they contain weaker results than those in the original paper. Therefore we would like to end with a commercial: Give Kermack's and McKendrick's paper a thorough try. It makes very rewarding reading!

The integral equations that result from our population models are always of the so-called Volterra type, *i.e.* the argument t of the unknown function b at the left hand side also delimits the integral occurring at the right. In our opinion there are two reasons for stressing the (re)formulation of models in terms of such integral equations. The first is entirely pragmatical: there are at present more results available in the mathematical literature for such integral equations than for the functional partial differential equations advocated in the previous chapters. The second reason is more of a philosophical nature. By writing down the integral equation direct from first principles, as was done in the last example, we circumvent the need to specify a detailed i-state space model, involving all sorts of possibly entirely unnecessary assumptions. Instead we concentrate on the minimal biological assumptions (1) and (2) from the previous subsection, thereby highlighting the range of applicability of our final results.

REMARK 1.3.5: The conditions allowing a population model to be replaced by an equivalent renewal equation are less restrictive than those allowing the construction of an age representation. In a constant environment the only condition is that individuals should pass through a stage in which they are all equal somewhere prior to giving birth. (For an age representation to apply this stage should occur immediately after birth.) For example, in the cell model from I.4 with $a \geqslant \frac{1}{2}$ (a was the size at which division could first occur) such a stage can be found in the reaching of size $\frac{1}{2}$, or any other fixed size between $\frac{1}{2}$ and a (compare II.10). The recruitment to the equalizing stage now replaces the birth rate function in equation (1.3.1).

1.4. The calculation of some birth kernels

In this subsection we shall illustrate the calculation of the birth kernel ϕ by means of some examples.

EXAMPLE 1.4.1: *Strictly age dependent reproduction and death* Assume that, provided an individual has not died yet, it gives birth at random, with birth events occurring in a Poisson process with rate $\lambda(a)$, where a denotes age, and that death occurs independent of an individual's history of giving birth, with a survivor function $\mathfrak{F}(a)$. In that case the birth kernel can be written as

$$\phi(a) = \mathfrak{F}(a)\,\lambda(a)\,. \tag{1.4.1}$$

The age distribution in this case also can be considered a p-state in the strict sense. The instantaneous death rate in a partial differential equation for the age distribution equals

$$\mu(a) = -\,d\log\mathfrak{F}(a)/da\,. \tag{1.4.2}$$

EXERCISE 1.4.2: Write down both the integral and the partial differential equation versions of this model assuming constant environments, and transform the latter into the former by integration along the characteristics. In passing you will also have derived a formula for the forcing function g in equation (1.3.1). How would you approach the problem of constructing the reverse transformation?

EXAMPLE 1.4.3: *Finite i-state spaces.* Assume that the i-state space is finite, and represented as $\Omega = \{1,..,k\}$. Condition (2′) from 1.2 and the definition of the state concept together imply that the state process is a so-called continuous time Markov chain with killing. Such a process can be characterized by its constant transition rates $b_{ij}, i \neq j$, where b_{ij} is the probability per unit of time that the chain will jump to i whenever it happens to be in j, and the death rate μ_j which we shall assume to be constant as well. Let $b_{jj} := -(\sum_i b_{ij} + \mu_j)$, *i.e.* $-b_{jj}$ is the rate at which j is left, and

$B := [b_{ij}]_{i,j=1,..,k}$. B is called the differential generator of the chain. The definition of the state concept also implies that during any period the chain is in state j birth events should occur in a so-called process with time invariant independent increments. We shall assume the birth process to be independent of external influences implying that the mean number of births per unit of time produced by an individual in state j is a constant which we shall call a_j. We shall denote the column vector* $(a_1,..,a_k)^T$ as A. Finally condition (1') from 1.2 compels us to assume that a neonate has a fixed probability c_j of being born in state j. We shall denote the column vector $(c_1,..,c_k)^T$ as C. Then

$$\phi(a) = A^T e^{Ba} C \tag{1.4.3}$$

and the probability that an individual survives to age a is

$$\mathfrak{P}(a) = E^T e^{Ba} C \tag{1.4.4}$$

where $E=(1,..,1)^T$.

EXERCISE 1.4.4: Derive (1.4.3) by writing a differential equation for the population state vector N in the form

$$\frac{dN}{dt} = (B + CA^T)N = BN + (A^T N)C \tag{1.4.5}$$

(NB according to the rules of matrix multiplication $A^T N$ is a scalar, so that $A^T NC = CA^T N$, while CA^T is a matrix of rank one), solving this equation as if $b := A^T N$ were known as a function of time, and substituting the solution in the defining formula for b. In passing you will have derived a formula for the forcing function g in equation (1.3.1).

EXERCISE 1.4.5: Calculate ϕ and $\mathfrak{P}$ for the special model depicted in figure 1.4.1 and plot the results.

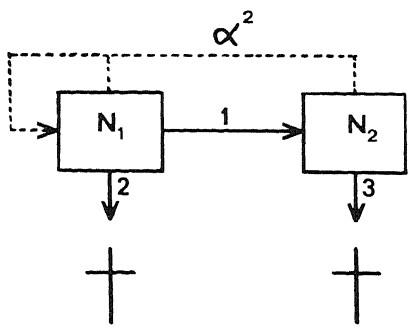

Fig. 1.4.1. Life cycle diagram corresponding to $B = \begin{bmatrix} -3 & 0 \\ 1 & -3 \end{bmatrix}$, $A^T=(0\ \ \alpha^2)$, $C=\binom{1}{0}$.

EXERCISE 1.4.6: Let the individual life cycle be given by the diagram of fig. 1.4.2 with α_1, α_2 and β constant. Assume that δ_1 and δ_2 each can take two values δ_i' and δ_i'' depending on the environment. First calculate the age dependent birth rate $\lambda(a) = \phi(a) / \mathfrak{P}(a)$ for a constant environment. Show that $\lambda(a)$ is the same for both environments if (* and only if) $\delta_1' - \delta_1'' = \delta_2'' - \delta_2''$. Also show that this condition on the death rates guarantees that the mean birth rate of a live individual is just a function of its age, no matter how the environment fluctuates. (Hint: Write down differential equations for the probabilities $p_j(t_0+a, a)$ that an individual born at t_0 is in state j at age a. Derive from this a differential equation for p_1+p_2. Finally show that $p_i / (p_1+p_2)$ satisfies a linear differential equation with constant coefficients.)

* ^{T} denotes transposition.

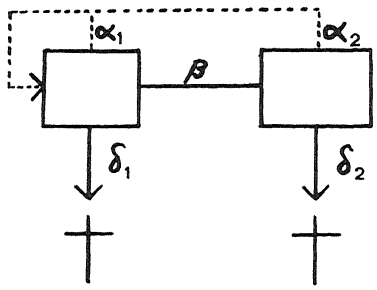

Fig. 1.4.2. Life cycle diagram corresponding to $B = \begin{bmatrix} -\beta+\delta_1 & 0 \\ \beta & \delta_2 \end{bmatrix}$, $A^T=(\alpha_1, \alpha_2)$ and $C = \binom{1}{0}$

REMARK 1.4.7: The main interest of example 1.4.3 is that the procedure outlined in exercise 1.4.4 can be reversed. Starting from a birth kernel like (1.4.3) it is always possible to derive an equivalent finite set of differential equations. This is especially helpful if we wish to do numerical studies on nonlinear problems. In section 5 we shall present this technique in more detail (and in section 2.3.3 we explore the connection with the representation theorems of linear systems theory).

EXAMPLE 1.4.8: *Deterministic binary fission combined with stochastic individual growth.* In III.1 and III.B we introduced a model in which an individual was assumed to grow according to a stochastic process of the diffusion type with size dependent infinitesimal mean and variance v and σ^2 respectively. Individuals were assumed to split into two equal parts on reaching size x_1, and to die on reaching size $x_0<\frac{1}{2}x_1$. If f_1 denotes the (defective*) probability density of the time it takes to grow from size $\frac{1}{2}x_1$ to size x_1 and f_0 the (defective) probability density of the time it takes to reach size x_0 starting from $\frac{1}{2}x_1$, then

$$\phi(a) = 2f_1(a) \tag{1.4.6}$$

and

$$\mathfrak{F}(a) = \int_a^\infty (f_0(\alpha) + f_1(\alpha))\, d\alpha\,. \tag{1.4.7}$$

(We assume here that there are no other causes of death. If the model also specifies a nonzero, possibly size dependent, death rate then we should add a third component in (1.4.7) denoting the (defective) probability density of the time till death from this cause.)
Calculating f_0 and f_1 is slightly complicated. In fact it is much easier to calculate their Laplace transforms

$$\bar{f}_i(s) := \int_0^\infty e^{-sa} f_i(a) da \tag{1..4.8}$$

As it turns out this is a bonus instead of a setback, as this Laplace transform is exactly the quantity we need in our calculations in section 2. We shall only give the procedure for $\bar{f}_1$.

Let $u_1(x,s)$ denote the Laplace transform of the (defective) probability density of the time an individual now sitting at x still needs to reach x_1, so that $\bar{f}_1(s)=u_1(\frac{1}{2}x_1,s)$. Then a well known result from the theory of diffusion processes (see e.g. COX & MILLER, 1965; GOEL & RICHTER-DYN, 1979; KARLIN & TAYLOR, 1981) tells us that u_1 solves the boundary value problem

$$\frac{1}{2}\sigma^2 \frac{d^2u_1}{dx^2} + v\frac{du_1}{dx} = su_1 \tag{1.4.9}$$

$$u_1(x_0,s) = 0, \quad u_1(x_1,s) = 1$$

In general we cannot solve (1.4.9) explicitly except for some special choices of v and σ^2. In 2.A we shall indicate how we can nevertheless use (1.4.9) to obtain useful bits of information about $\bar{f}_1$.

---

* A probability density f is called defective if $\int f(t)dt<1$.

EXERCISE 1.4.9: Indicate how for a given initial size distribution we might calculate the Laplace transform of the forcing function g in equation 1.3.1.

2. Linear theory

We shall assume throughout this section that the physical environment is constant and that individuals interact neither directly nor indirectly. All the complexities of the (stochastic) i-behaviour then can be summarized by an age dependent mean birth rate $\lambda(a)$ and survival probability $\pi(a)$, or even more concisely by the birth kernel $\phi(a)=\lambda(a)\pi(a)$. We start with equation (1.3.1). In subsection 2.3 we shall also consider for the first time a p-equation for the age distribution.

2.1. An explicit expression for the population birth rate

To solve (1.3.1) we again use the biological interpretation. But first we introduce some notation to simplify our chores.

Given two locally integrable functions g_1 and g_2 on $\mathbb{R}^+$ we define their *convolution*, written as g_1*g_2 to be the function*

$$(g_1*g_2)(t) := \int_0^t g_1(t-\tau)g_2(\tau)d\tau\,. \tag{2.1.1}$$

The following relations hold (see also remark 2.1.4 for the precise interpretation of the equal sign)

$$g*(\lambda g_1 + \mu g_2) = \lambda g*g_1 + \mu g*g_2\,, \tag{2.1.2}$$

$$g_1*g_2 = g_2*g_1\,, \tag{2.1.3}$$

$$\delta*g = g\,, \tag{2.1.4}$$

where $\lambda,\mu\in\mathbb{R}$ and δ is the delta "function" introduced in III.5 (if you have skipped III.5 you may consider (2.1.4) as just the definition of δ). That is, functions under convolution behave pretty much like ordinary numbers under multiplication, with δ playing the role of 1. This is brought out even more sharply if we look at their Laplace transforms

$$\bar{g}(s) := \int_0^\infty e^{-st}\, g(t)dt\,, \tag{2.1.5}$$

for

$$\overline{g_1*g_2}(s) = \bar{g}_1(s)\,\bar{g}_2(s)\,. \tag{2.1.6}$$

(Some general information about the Laplace transform for non-cognoscenti may be found in the appendix to this chapter.)

EXERCISE 2.1: Verify (2.1.2) to (2.1.4) and also (2.1.6). If necessary assume that g is continuous.

Using (2.1.1) we can write (1.3.1) as

$$b = \phi*b + g = b*\phi + g\,, \tag{2.1.7}$$

where, of course, we make the biologically natural assumption that both ϕ and g are nonnegative as well as locally integrable.

Now imagine an individual born at $t = 0$. By the same argument by which we derived (1.3.1) we can calculate the (mean) rates at which the various generations of its descendants are born as

$$\begin{aligned}
&1^{st}\text{ generation births: } && \phi^{*1} := \phi\\
&2^{nd}\text{ generation births: } && \phi^{*2} := \phi*\phi^{*1}\\
&\quad\vdots && \quad\vdots\\
&n^{th}\text{ generation births: } && \phi^{*n} := \phi*\phi^{*(n-1)}
\end{aligned} \tag{2.1.8}$$

* The Fubini Theorem implies that g_1*g_2 is a well-defined locally integrable function; see RUDIN (1974) 7.13.

Therefore the total (mean) birth rate of its clan as a function of time equals

$$\Phi := \sum_{n=1}^{\infty} \phi^{*n} \tag{2.1.9}$$

An elegant proof of the convergence of this sum may be found in JAGERS (1975) p. 105. Also see DOETSCH (1956) III.25. In the literature on Volterra integral equations *(e.g.* MILLER, 1971) Φ is known as the resolvent of (2.1.7), and the equation $\Phi = \phi*\Phi+\phi$ is called the resolvent equation.

Remembering that g was the birth rate from individuals already present before $t=0$ and from the outside world, we infer that

$$b = g + \Phi*g\ , \tag{2.1.10}$$

i.e. the total birth rate equals the starter birth rate g plus the cumulation of all the birth rates from clans started by it. (2.1.10) is an example of a so-called generation expansion, already referred to in I.5 (another example may be found in II.3).

EXERCISE 2.1.2: Verify that (2.1.10) satisfies (2.1.7)

EXERCISE 2.1.3: Simplify (2.1.10) by adapting the definition of Φ, using (2.1.4).

Since we derived (2.1.10) by the same reasoning by which we derived (2.1.7) anyway, there is not much sense in proving that (2.1.10) is the only solution: it is the only solution that matters. However, if we wish to use (2.1.7) as a basis for other calculations we have to make sure that the solution is unique. A somewhat formal proof goes as follows. Assume that b_1 and b_2 are two, possibly different, solutions. Then

$$b_1 - b_2 = (b_1-b_2)*\phi$$

and consequently, by induction,

$$b_1 - b_2 = (b_1-b_2)*\phi^{*n} \to 0$$

(for otherwise the sum in (2.1.9) would not converge), *i.e.* $b_1=b_2$. This proof can be made rigorous in various ways (*i.e.*, under various assumptions on g).

(* Alternatively, let $\mathbf{1}$ denote the function which is identically one, then $\mathbf{1}*((b_1-b_2) = \mathbf{1}*\phi*(b_1-b_2)$ and consequently $(\mathbf{1}-\mathbf{1}*\phi)*(b_1-b_2) = 0$. The Titchmarsh Theorem (Doetsch (1950) I.2.15 Satz 11 and 12) implies that necessarily either $b_1-b_2 \equiv 0$ or $\mathbf{1}-\mathbf{1}*\phi \equiv 0$. The latter is impossible since $\mathbf{1}*\phi(t) \to 0$ for $t \downarrow 0$).

* REMARK 2.1.4: Some care is needed in interpreting the various equal signs. In general we only have equality almost everywhere. This is entirely sufficient for practical purposes. If g is continuous, then we can decide to restrict ourselves to continuous solutions only, in which case the equal signs in (2.1.7), (2.1.10) and that in $b_1=b_2$ may be interpreted as referring to pointwise equality.
Formulae (2.1.9) and (2.1.10) provide an adequate means to calculate b for small t, simply by truncating (2.1.9) after a few terms. For large t the representation (2.1.9) has no practical value whatsoever. However, the large time behaviour is covered by a neat set of mathematical theorems reviewed in the next subsection. (But there is a price to pay: see subsection 2.3.)

2.2. *Renewal theorems*

We start with some heuristics. From previous experience we guess that b will eventually grow about exponentially. However, in contrast to the differential equation case dealt with in the previous chapters we cannot simply proceed by substituting a trial solution of the suspected form in (1.3.1) in order to find the possible growth rates supported by the dynamics, as in the integral equation case the representation of the dynamics comes in one package with the initial condition in the form of g. Apparently we have to get rid of the initial condition first. We can do this by considering an equation analogous to (1.3.1) but living on the whole time axis instead of just the positive half axis:

$$b(t) = \int_0^{\infty} b(t-a)\phi(a)da\ . \tag{2.2.1}$$

REMARK 2.2.1: We can derive (2.2.1) formally by substituting $b'(t) := b(t+t_0) / b(t_0)$ *(i.e.* we shift the time origin to

$-t_0$ and renormalize to $b'(0)=1$) in (1.3.1) and letting $t_0 \to \infty$. This indicates that information gleaned from (2.2.1) about the possible asymptotic behaviour of b can only be relevant in cases where $g(t)/b(t) \to 0$.

Substituting a trial solution of the form $b(t)=e^{rt}$ in (2.2.1) we find that r should satisfy

$$1 = \bar{\phi}(r) \tag{2.2.2a}$$

where

$$\bar{\phi}(s) := \int_0^\infty e^{-sa}\, \phi(a)da \tag{2.2.2b}$$

is the Laplace transform of the birth kernel ϕ.

Equation (2.2.2a) is known as the *characteristic equation*. We can without loss of biological generality assume that ϕ increases at most exponentially, so that there exists a $\sigma \geqslant -\infty$ such that the integral in (2.2.2b) converges absolutely for $\mathrm{Re}(s)>\sigma$ and not for $\mathrm{Re}(s)<\sigma$. If we wish r to qualify as the asymptotic rate of exponential growth of our birth function b, then r should be real. So we concentrate on the behaviour of $\bar{\phi}$ for real s. Since $\phi(a) \geqslant 0$ the Laplace transform $\bar{\phi}$ is a monotonically decreasing smooth function of s with $\bar{\phi}(+\infty)=0$, and with $\lim_{s \downarrow \sigma} \bar{\phi}(s) = \theta \leqslant \infty$ (if $\sigma=-\infty$, which is the case when $\phi(a)$ goes to zero sufficiently fast, then necessarily $\theta=+\infty$). If $\theta<1$ then (2.2.2) has no solution. Therefore we shall for the time being assume that $\theta \geqslant 1$. In that case (2.2.2) allows a unique solution. This solution r is called the *intrinsic rate of natural increase* associated with the birth kernel ϕ.

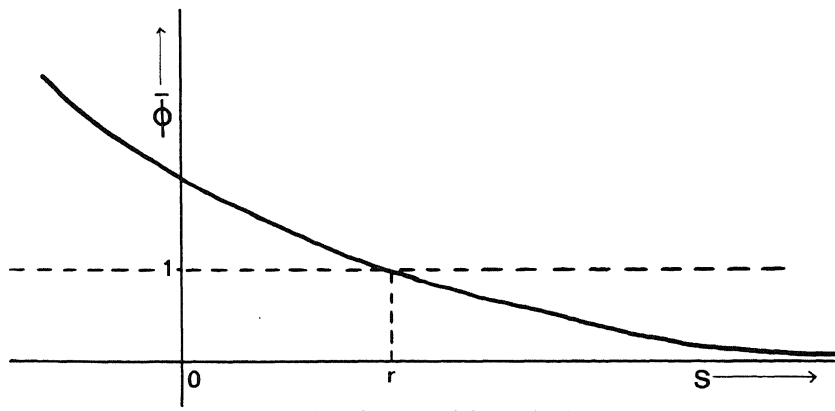

Fig. 2.2.1: Solving the characteristic equation for r

Some information about the whereabouts of r can be obtained from the observation that (see fig. 2.2.1)

$$\bar{\phi}(0) \gtrless 1 \Leftrightarrow r \gtrless 0 \tag{2.2.3a}$$

in combination with

$$\bar{\phi}(0) = R\,, \tag{2.2.3b}$$

where R is the so-called *net reproductive number, i.e.* the mean number of offspring a newborn individual is expected to beget during her lifetime. In general r has to be calculated numerically. Appendix 2A provides some convenient approximations.

EXERCISE 2.2.1: Check both (2.2.3a) and (2.2.3b).

EXERCISE 2.2.2: What does the characteristic equation from example 1.4.3 look like? What do you find for the characteristic equation of the linear differential equation (1.4.5)? Calculate r explicitly for the special case of exercise 1.4.5.

EXERCISE 2.2.3: Calculate R for example 1.4.3. Hint: Use $R = \bar{\phi}(0)$ and $\exp(Ba)=\frac{d}{da} B^{-1}\exp(Ba)$.

EXERCISE 2.2.4: Calculate R for example 1.4.8. Hint: Put $s=0$ in equation (1.4.9).

The statement that $b(t)$ will grow (or decrease) eventually exponentially at a relative rate r is the content of the so-called renewal theorem. This theorem comes in various disguises depending on what kind of assumptions one makes concerning ϕ and g. In a population context estimates for the integral of b are quite appropriate (indeed only the integral is a "number"), therefore we first present a theorem in this spirit and only thereafter give a variant which yields a much stronger conclusion, but at the expense of a much stronger assumption about g. But first of all we formulate the main technical tool as

LEMMA 2.2.5: *Assume that for some* $\delta>0$ *(i)* $\int_0^\infty e^{-(r-\delta)\tau}\phi(\tau)d\tau<\infty$ and *(ii)* $\bar{\phi}(s)\neq 1$ *for* $r-\delta\leq \mathrm{Re}s<r$, *then*

$$b(t) = \frac{e^{rt}\int_0^t e^{-r\tau}g(\tau)d\tau}{\int_0^\infty \tau e^{-r\tau}\phi(\tau)d\tau} + g(t) + \tilde{\Phi}*g(t) \tag{2.2.4}$$

where $\tilde{\Phi}$ *is such that* $\int_0^\infty e^{-(r-\delta)\tau}|\tilde{\Phi}(\tau)|d\tau < \infty$.

Note that the possible choices for δ in lemma 2.2.5 are restricted by two conditions. (i) refers essentially to the behaviour of ϕ at infinity; a sufficient condition for (i) is that $\phi(t)<ce^{(r-\delta-\epsilon)t}$ for all $t>T$, for some positive T and ϵ. (ii) refers to the position of any "other" roots of the characteristic equation in the complex plane; it can be shown that, for nonnegative ϕ, such roots always lie somewhat to the left of a the line $\mathrm{Re}s=r$, so there always exists a $\delta>0$ satisfying (ii). The next remark gives references for the proof of the lemma.

* REMARKS 2.2.6: (i) Lemma 2.2.5 is a generalization of a famous theorem of Paley and Wiener stating that the resolvent Φ is in L_1 whenever all roots of the characteristic equation lie in the left half plane (see *e.g.* MILLER, 1971). In lemma 2.2.5 the resolvent Φ is represented as $\Phi_0 + \tilde{\Phi}$, where $\Phi_0(t) = (\int_0^\infty \tau e^{-r\tau}\phi(\tau)d\tau)^{-1}e^{rt}$ corresponds to the pole of $\bar{\Phi}$ at r, and $\tilde{\Phi}$ lies in a weighted L_1 space. A proof that such a representation is possible may be found in JORDAN & WHEELER (1980a, b) or JORDAN, STAFFANS & WHEELER (1982).
(ii) In DIEKMANN & VAN GILS (1984) one can find indicated how lemma 2.2.5 fits in with the semi-group approaches sketched in subsection 2.3.2.
(iii) Of course one can generalize the representation (2.2.4) by calculating residues in additional poles of $\bar{\Phi}$.

The following two "renewal theorems" are direct consequences of lemma 2.2.5.

THEOREM 2.2.7: *If, in addition to the assumptions on* ϕ *from lemma 2.2.5,* $\int_0^\infty e^{-(r-\delta)\tau}g(\tau)d\tau<\infty$, *then*

$$b(t) = e^{rt}\frac{\int_0^\infty e^{-r\tau}g(\tau)d\tau}{\int_0^\infty \tau e^{-r\tau}\phi(\tau)d\tau} + f(t)\,, \tag{2.2.5}$$

where f *satisfies* $\int_0^\infty e^{-(r-\delta)\tau}|f(\tau)|\,d\tau<\infty$, *and consequently for any weighting function* h *with* $e^{-(r-\delta)a}h(a)$ *both bounded and integrable*

$$\int_0^t b(t-a)h(a)da = e^{rt}\frac{\int_0^\infty e^{-r\tau}g(\tau)d\tau \int_0^\infty e^{-ra}h(a)da}{\int_0^\infty \tau e^{-r\tau}\phi(\tau)d\tau} + 0(e^{(r-\delta)t})$$

THEOREM 2.2.8: *If, in addition to the assumptions on* ϕ *from lemma 2.2.5* $e^{-(r-\delta)t}g(t)$ *is bounded, then*

$$b(t) = e^{rt}\frac{\int_0^\infty e^{-r\tau}g(\tau)d\tau}{\int_0^\infty \tau e^{-r\tau}\phi(\tau)d\tau} + 0(e^{(r-\delta)t}) \tag{2.2.7}$$

REMARKS 2.2.9: (i) One can deal with systems of renewal equations in much the same manner as we did with one.
(ii) We have only used the nonnegativity of ϕ to guarantee the existence of a strictly dominant real root r. Appropriately modified forms of lemma 2.2.5 and theorems 2.2.7 and 2.2.8 apply when we dispense with the assumptions about the sign of ϕ and g.

In lemma 2.2.5 and theorems 2.2.7 and 2.2.8 we have concentrated on exponential estimates for the remainder term under the crucial hypothesis that for some positive δ the assumptions of lemma 2.2.5 hold. The following result which is essentially due to Feller and which with some right is referred to as *the* renewal theorem, makes no such hypothesis but has, as a consequence, a much weaker conclusion about the remainder term.

THEOREM 2.2.10: *Suppose that (i) g is continuous almost everywhere and (ii)* $\sum_{k=0}^{\infty} \sup_{0\leq\tau<1} e^{-rk} g(k+\tau) < \infty$ *then*

$$e^{-rt} b(t) \to \frac{\int_0^\infty e^{-r\tau} g(\tau) d\tau}{\int_0^\infty \tau\, e^{-r\tau}\, \phi(\tau) d\tau} \quad \text{for } t \to \infty\,, \tag{2.2.8}$$

where the right hand side is interpreted as zero when the integral in the denominator diverges.

(A sufficient condition for (i) is that g is piecewise continuous and a sufficient condition for (ii) is that $e^{-rt} g(t) < f(t)$ with f decreasing and integrable.) A proof of this theorem may be found in JAGERS (1975).

REMARK 2.2.11: The renewal theorem essentially relates the asymptotic behaviour of the solution of the renewal equation (1.3.1) to the positive solutions of the "limiting equation" (2.2.1). It is this relationship which will be extended into the nonlinear realm in section 4.

We finish with two theorems which also apply when $\bar{\phi}(\sigma) < 1$. The first theorem provides some general insight in the behaviour of b when $R < 1$, the second one provides a general estimate for b. Proofs may be found respectively in JAGERS (1975) and HOPPENSTEADT (1975).

THEOREM 2.2.12: *If $R < 1$ and $\lim_{t\to\infty} g(t) = \gamma$ then*

$$\lim_{t\to\infty} b(t) = \gamma / (1-R) \tag{2.2.9}$$

COROLLARY 2.2.13: *If $R < 1$ and $g \to 0$ then $b \to 0$.*

EXERCISE 2.2.14: Show that when $R < 1$

$$\int_0^\infty \Phi(\tau) d\tau = R / (1-R) \tag{2.2.10}$$

and use this observation to interpret (2.2.9).

THEOREM 2.2.15: *For all $\lambda \in \mathbb{R}$ such that $\bar{\phi}(\lambda) < 1$*

$$e^{-\lambda t} b(t) \leq \frac{\sup_{0\leq\tau<\infty} g(\tau) e^{-\lambda\tau}}{1-\bar{\phi}(\lambda)}\,. \tag{2.2.11}$$

EXERCISE 2.2.16: Show that $\tilde{b}(t) := e^{-\lambda t} b(t)$ satisfies a renewal equation just like b, with $\tilde{\phi}(a) := e^{-\lambda a}\phi(a)$ and $\tilde{g}(t) := e^{-\lambda t} g(t)$.

* EXERCISE 2.2.17: Use the result of the previous exercise to prove theorem 2.2.15.

* EXERCISE 2.2.18: Derive the analogue of (2.2.11) for $\lambda \in \mathbb{C}$. (Hint: Take moduli where convenient.) Use this result to conclude that under the conditions of theorem 2.2.8 the Laplace transform of b exists for $\mathrm{Re}(s) > r$.

* EXERCISE 2.2.19: Use the complex inversion formula (A.13) together with a shift of contour as indicated in figure 2.2.2. to prove theorem 2.2.8.
Hint: Show that both $\bar{g}(s)$ and $\bar{\phi}(s) \to 0$ on the segments pp' and gg' for $\mathrm{Im}(p) = -\mathrm{Im}(q) \to \infty$ and that therefore $\bar{b}(s) \to 0$ on these segments. Therefore we can calculate b as the limit for $\mathrm{Im}(p) \to \infty$ of $\frac{1}{2\pi i} \int_q^p e^{st} \bar{b}(s) ds$ plus the sum of the residues of $\bar{b}(s) e^{st}$ within the contour. Show that the former term is $0(e^{(r-\delta)t})$.

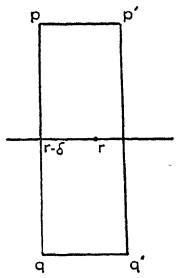

Fig. 2.2.2. Shift of contour used in exercise 2.2.19

2.3. *Semigroup approaches*

The renewal theorem provides an easy, cut and dried tool for most of our practical needs. The cost we pay for exclusively relying on it is that we lose contact with the semigroup and dynamical systems framework as expounded in chapter II and thereby with our easiest route towards nonlinear extensions, stability theory, bifurcation theory, *etc.*, as discussed in chapter VI. In this section we shall make up for this deficiency. In the spirit of chapter II we shall restrict ourselves throughout to the autonomous case, *i.e.* we shall assume that there are no births from outside sources. Moreover our approach will be almost entirely heuristic. Our aim is to sketch the various approaches, theories and results, and in particular their interrelationships. We shall not be very precise in the mathematical formulation, neither do we try to present the most general or sharpest results. More complete treatments may be found in WEBB (1984, 1985a) and in DIEKMANN (1980), VAN GILS (1984) and STAFFANS (1984).

2.3.1. *The age distribution*

The traditional biological way to define a dynamical system is in terms of the age density n, where* $n \in L_1(0, a_m)$, $a_m \leqslant \infty$, say. The semigroup transforming age distributions into each other will be called T_2. (The reason for the choice of index will become clear below.) Let the mean number of offspring produced per unit of time by an individual aged a be denoted as $\lambda(a)$, and the probability that individuals survive to age a as $\Pi(a)$. We shall assume that λ is (essentially) bounded. To calculate T_2 we observe that an animal aged a at $t=0$ has a probability $\Pi(a+t)/\Pi(a)$ to survive to time t and age $a+t$. Moreover, at time t an individual of age $a<t$ was born into the population at time $t-a$, and of the individuals born at that time only a fraction $\Pi(a)$ are still alive Therefore†

$$n(t,a) = \begin{cases} n(0,a-t)\Pi(a)/\Pi(a-t) & \text{for } a \geqslant t \\ b(t-a)\,\Pi(a) & \text{for } a < t \end{cases} \tag{2.3.1a}$$

where b is calculated from (2.1.10) with

* See chapter II section 1 or III.A.2 for a description of the function space L_1.

† We shall write $n(t,a)$ for $n(t)(a)$, the age density at t evaluated at a. Moreover, we shall often suppress either of the arguments depending on the context.

$$\phi(a) = \lambda(a)\mathfrak{F}(a), \qquad g(t) = \int_0^{a_m} n(0,a)\,\lambda(a+t)\,\frac{\mathfrak{F}(a+t)}{\mathfrak{F}(a)}\,da\,. \tag{2.3.1b}$$

In the spirit of chapter II the asymptotic behaviour of T_2 should be deduced from the consideration of its differential generator A_2. For the full details of how this can be done you are referred to WEBB (1984, 1985a). We shall confine ourselves to deriving A_2 and calculating its dominant eigenvalue r.

To calculate A_2 we observe that $n(t+dt,\ a+dt)=n(t,a)\ \mathfrak{F}(a+dt)\,/\,\mathfrak{F}(a)$. Subtracting $n(t,\ a+dt)$ from both sides and writing $\mathfrak{F}(a+dt)\,/\,\mathfrak{F}(a)=1-\mu(a)dt$ with $\mu=-d\log\mathfrak{F}\,/\,da$, gives after division by dt

$$\frac{\partial n}{\partial t} = -\frac{\partial n}{\partial a} - \mu(a)n \tag{2.3.2}$$

and therefore

$$A_2 = -\frac{\partial}{\partial a} - \mu \tag{2.3.3a}$$

provided* $n \in AC$. Moreover $n(0)$ should equal b so the domain of A_2 is given by

$$\mathfrak{D}(A_2) = \{n \mid n \in AC\ [0,a_m] \text{ and } n(0) = \int_0^{a_m} \lambda(a)n(a)da\} \tag{2.3.3b}$$

EXERCISE 2.3.1: Derive the explicit formulae for λ and μ from the model of fig. 1.4.1.

The first to derive the partial differential equation (2.3.2) was MCKENDRICK (1926). Later it was rediscovered by VON FOERSTER (1959) and many others. The explicit semigroup interpretation appeared in the literature only relatively recently in the work of WEBB (1979, 1981, 1983a,b, 1984, 1985a) and PRÜSS (1981, 1983a,b).

To calculate r we write $A_2\tilde{n}=r\tilde{n}$ to find

$$\tilde{n}(a) = \tilde{n}(0)e^{-ra}\,\mathfrak{F}(a)\,. \tag{2.3.4}$$

The condition that $\tilde{n}$ should lie in $\mathfrak{D}(A_2)$ then gives

$$1 = \int_0^{a_m} \lambda(a)e^{-ra}\,\mathfrak{F}(a)da = \int_0^{\infty} e^{-ra}\phi(a)da \tag{2.3.5}$$

which is just our old acquaintance (2.2.2). Along the lines expounded in chapter II one can prove (WEBB 1984, 1985a) that under appropriate conditions

$$[T_2(t)n_0](a) = e^{rt}c(n_0)\mathfrak{F}(a)e^{-ra} + 0(e^{(r-\epsilon)t}) \text{ for } t\to\infty\,, \tag{2.3.6}$$

where n_0 is the initial age density and $c(n_0)$ is a constant depending linearly on n_0. In appendix 2.B we shall discuss how c can be calculated direct from the interpretation.

EXERCISE 2.3.2: Derive (2.3.6) from the renewal theorem and calculate $c(n_0)$

* REMARK 2.3.3: When $a_m<\infty$ then it is easy to show that the linear (sub)space $F=\{n\in L_1(0,a_m) \mid \|n\|_B<\infty\}$ with $\|n\|_B = \int_0^{a_m} |\,n(a)\,|\,/\,\mathfrak{F}(a)da$ is invariant under T_2. Moreover $L_1(0,a_m)$ gets mapped into F by any $T_2(t)$ with $t\geq a_m$. Therefore we may just as well choose F provided with the norm $\|\cdot\|_B$ instead of $L_1(0,a_m)$ as the space on which T_2 is supposed to act. As an alternative equivalent we may, on the analogy of the procedure followed in chapter II, make a change of variable $m(t,a)=n(t,a)\,/\,\mathfrak{F}(a)$, with m assumed to be in $L_1(0,a_m)$. In this manner the death rate μ is eliminated from the differential generator.

---

* AC denotes the space of absolutely continuous functions, *i.e.* functions n such that $\int^a n(\alpha)d\alpha \in L_1$. See also chapter II definition 2.8 for an alternative characterization. The differential operator in (2.3.3a) should be interpreted as just the (left)inverse of the integration operator.

2.3.2. *Two semi-groups derived directly from the renewal equation itself*

There exist also less traditional ways to define semigroups for age dependent processes. Following DIEKMANN (1980) we shall discuss here two such semigroups based directly on two variants of the renewal equation.

The initial data enter (1.3.1) in the form of the function g. This therefore seems the obvious candidate as a basis for defining a semigroup. The question then is what becomes of g if we write down an equation for b from $\tau>0$ onwards and shift the time origin to τ. In the following we shall append an index τ to g to express the shift of time origin, g_0 being just our original g. From the dynamical viewpoint g_τ is the state of the system at time τ. Either by manipulating (1.3.1) or direct from the interpretation we find

$$g_\tau(t) = g_0(t+\tau) + \int_0^\tau b(s)\,\phi(t+\tau-s)ds \tag{2.3.7}$$

with b calculated from (2.1.10). Rewriting this in dynamical systems notation gives

$$T_1(\tau)g_0 = g_\tau \tag{2.3.8}$$

which may be considered the definition of the semigroup T_1. To complete this definition we still have to indicate which underlying function space for the "forcing" functions g we are dealing with. We shall postpone discussing this problem till after the introduction of yet another semigroup.

EXERCISE 2.3.4: Calculate the differential generator A_1 of T_1. Compare the result with that of exercises I.2.3.4 to 2.3.8. Calculate the characteristic equation corresponding to A_1 and make a "guess" at the asymptotic behaviour of $T_1(\tau)g_0$.

Usually the only sure way to know an animal's age is to observe when it was born. However, if our data include the history of the birth events then clearly there is no need to go through the intermediate stage of calculating an age distribution if all we wish to calculate anyhow are the future values of b. This can just as well be done from equation (2.2.1):

$$b(t) = \int_0^\infty \phi(a)b(t-a)da$$

The initial data for this equation are the values of b for $t\leqslant 0$. For later use we shall put $b(t)=:h_0(t)$ for $t\leqslant 0$, and rewrite (2.2.1) as

$$b(t) = \int_0^t \phi(a)b(t-a)da + \int_t^\infty \phi(a)h_0(t-a)da \tag{2.3.9}$$

Using the analogy with the previous case we can immediately rewrite this as the semigroup

$$(T_3(\tau)h_0)(t) = h_\tau(t) := b(t+\tau),\quad t\leqslant 0,\ \tau\geqslant 0\,, \tag{2.3.10}$$

where $b(t)=h_0(t)$ for $t\leqslant 0$ and for $t>0$ can be calculated from (2.1.10) with

$$g(t) = \int_t^\infty \phi(a)h_0(t-a)da\,. \tag{2.3.11}$$

EXERCISE 2.3.5: Calculate the differential generator A_3 of T_3. Compare the result with that of exercise I.2.3.9. Calculate also the characteristic equation corresponding to A_3, and make a "guess" at the asymptotic behaviour of $T_3(\tau)h_0$.

The (most obvious) relation between the semigroups T_1 to T_3 is indicated in the diagram in fig. 2.3.1 showing some of the relevant mappings between the, as yet unspecified, spaces on which they act. Here S_i are so-called intertwining maps which means that *e.g.* $S_1T_3(\tau)h_0 = T_2(\tau)S_1h_0$ *etc.*

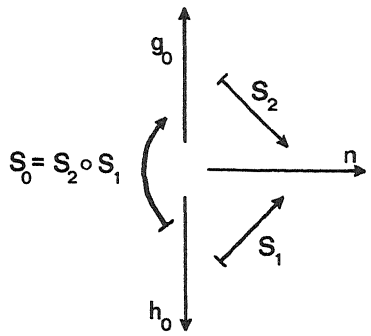

Fig. 2.3.1. Some maps connecting the semigroups T_1 to T_3.

The map S_0 is given by formula (2.3.11), showing that except for a choice of the spaces on which T_1 and T_3 act it is possible to reconstruct T_1 and T_3 from S_0 (provided H is sufficiently "rich" so that we may calculate the kernel ϕ from the action of S_0 on elements of H).

EXERCISE 2.3.6: Calculate S_1 and S_2 and discuss the degree to which the relation between T_1 to T_3 is unique.

As a last step in our definitions we still have to choose the spaces G and H on which T_1 and T_3 are supposed to act. In general this choice will depend on our practical needs. One possibility is to choose just $L_1(\mathbb{R}^-)$ for H and for G the closure of the image of H under the intertwining map S_0 (observe that $S_1 H$ is just equal to F from remark 2.3.3 and therefore also closure$(S_2 F) = G$). But this certainly is not the only possible choice.

The semigroup T_1 is studied in detail in DIEKMANN (1980) under the assumption that the kernel ϕ has compact support. One of the main results is that, with an appropriate choice of spaces, T_1 and T_3 are adjoints of each other. Many additional results are presented in VAN GILS (1984) (also see DIEKMANN (1982) and VERDUYN-LUNEL (1984, to appear)). STAFFANS (1984) gives a particularly nice overview covering intertwining as well as adjointness relations for both the cases of compact support, or finite delay, and infinite delay.

The semigroups T_1 to T_3 may be used as the basis for various nonlinear extensions. As none of them is based on a detailed conception of mechanism the choice between them is mainly a matter of convenience. A direct scientific argument for such a choice could be how the data are collected. Obviously T_3 and T_2 score best in this respect. From the mathematical point of view T_1 is especially pleasant to work with. However, the biological conditions allowing a problem to be formulated as a nonlinear extension of T_1 appear to be rather restrictive. Some examples may be found in section 4 and chapter VI (see also DIEKMANN & VAN GILS (1984)). T_2 certainly allows the largest scope for nonlinear extensions, the only (still rather strong!) restriction being that i-behaviour should fullfil conditions (1) and (2) from subsection 1.2. Mathematical references are PRÜSS (1983a,b), SCHAPPACHER (pers.comm.) and especially WEBB (1985a).

* 2.3.3. *Finite representability*

As a close to this subsection we shall say a few words about the possibility of representing our birth dynamics by means of a finite set of differential equations, as in the models covered by example 1.4.3 (see in particular exercise 1.4.4).

In the example we also have available a (semi)group T_4 acting on $\mathbb{R}^k$ obtained from solving the differential equation (1.4.5):

$$T_4(t) = \exp[(B + CA^T)t], \tag{2.3.12}$$

B a square matrix, A and C single column matrices. Adding this setting to our intertwining diagram gives fig. 2.3.2,

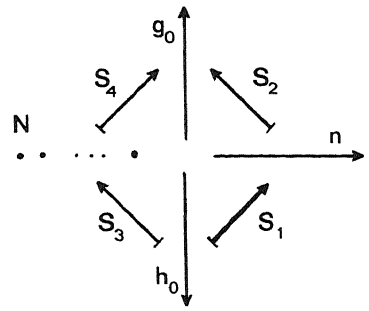

Fig. 2.3.2. The intertwining of the semigroups T_1 to T_4.

with

$$N = S_3 h_0 = \int_0^t e^{Ba} C h_0(-a) da \,. \tag{2.3.13}$$

and

$$g_0(t) = (S_4 N)(t) = A^T e^{Bt} N \,, \tag{2.3.14}$$

while the map $S_0 = S_4 \circ S_3 = S_2 \circ S_1$ can, as always, be written as

$$S_0 h_0 = \int_t^\infty \phi(a) h_0(t-a) da \tag{2.3.15a}$$

where now

$$\phi(a) = A^T e^{Ba} C \,. \tag{2.3.15b}$$

From the diagram it is clear that $S_0 = S_2 \circ S_1 = S_4 \circ S_3$ has a range of dimension $k' \leq k$. (2.3.15) moreover allows us to deduce that generically $k' = k$.

Given a linear map $S_0 : H \to G$ which has k-dimensional range it is always possible to factorize this map into $S_0 = S_4 \circ S_3$ where $S_3 : H \to \mathbb{R}^k$ and $S_4 : \mathbb{R}^k \to G$. If this map is moreover of the form (2.3.15a) and H is sufficiently rich (*e.g.* $L_1(\mathbb{R}^-)$ will certainly do) then the results from systems theory* tell us that it is possible to calculate from S_0 a $k \times k$ matrix B and $k \times 1$ matrices A and C, such that (2.3.15b) holds. A, B and C are unique up to a choice of basis for the "intervening" space.

The previous results can be rephrased as: If (and only if) S_0 has a finite dimensional range it can be factorized into $S_0 = S_4 \circ S_3$ with $S_3 : H \to \mathbb{R}^k$, $k = dim(S_0 H)$. If such is the case S_3 induces a semigroup T_4 of the form (2.3.12) on $\mathbb{R}^k$ which intertwines with T_3 and T_1 in the manner of figure 2.3.2 with S_4 an isomorphism between T_4 and the restriction of T_1 to $S_0 H$. The factorization and therefore also T_4 are unique up to a basis change. We shall refer to T_4 as a finite dimensional representation of our birth kernel ϕ.

From (2.3.15b) it can be seen immediately that a birth kernel allows a finite dimensional representation if and only if it can be written as a finite linear combination of polynomials times complex exponentials.

In general there is no guarantee that the components of the matrices A, B and C derived from a finite dimensionally representable kernel ϕ can be given signs compatible with a mechanistic model of the type described in example 1.4.3. Characterizing kernels which allow a mechanistically interpretable finite dimensional representation still seems to be an open problem.

In subsection 5.2 we shall consider the extension of the finite dimensional representability problem to the nonlinear case.

EXERCISE 2.3.7: Let $\phi(a) = pe^{-\alpha_1 a}$, $p, \alpha > 0$. Calculate the map S_0 and show that it has finite dimensional range. Find

* A survey of these results with an eye towards animal behaviour may be found in METZ (1981); a more readily obtainable reference is KALMAN, FALB & ARBIB (1969).

a differential equation model generating the same birth process.

EXERCISE 2.3.8: Let $\phi(a)=p_1e^{-\alpha_1 a}+p_2e^{-\alpha_2 a}$, $p_1, p_2, \alpha_1, \alpha_2>0$. Same questions as exercise 2.3.7.

EXERCISE 2.3.9: Let $\phi(a)=pae^{-\alpha a}$, $p,\alpha>0$. Same questions as exercise 2.3.7. Hint: remember exercise 1.4.5.

EXERCISE 2.3.10: Let $\phi(a)=p_1e^{-\alpha_1 a}+p_2e^{-\alpha_2 a}$, $\alpha_1>\alpha_2$, $p_2>-p_1>0$. Same questions as exercise 2.3.7. Also try to find a mechanistically interpretable representation.

EXERCISE 2.3.11: What are all possible ϕ you can get from two differential equation representations of the birth law, without violating the constraint that ϕ be nonnegative? Do all these ϕ allow a mechanistically interpretable representation as well?

* EXERCISE 2.3.12: Let $\phi(a)=pe^{-\alpha a}+qe^{-\beta a}\cos(\omega a+\theta)$, $\beta>\alpha$, $p>q>0$. Show that S_0 has a three dimensional range. Find a (not necessarily interpretable) differential equation representation of the corresponding birth process.

** EXERCISE 2.3.13: Can you find an interpretable representation of the ϕ from the previous exercise as well?

* EXERCISE 2.3.14: Let

$$\phi(a) = \sum_{i=1}^{h}\sum_{j=0}^{k_i} p_{ij}\, a^j e^{-\alpha_i a} + \sum_{i=1}^{m}\sum_{j=1}^{n_i} q_{ij}\, a^j\, e^{-\beta_i a}\cos(\omega_i a+\theta_{ij})\,.$$

What is dim range S_0? Could you construct a (not necessarily interpretable) differential equation representation of the corresponding birth process?

** EXERCISE 2.3.15: Characterize all birth kernels ϕ allowing a mechanistically interpretable representation by finitely many differential equations.
Hint: The problem has some analogy with the problem of representing stochastic processes as functions of Markov chains (see e.g. ERICKSON, 1970).

2.A. Moments, cumulants and some approximations for r

In this appendix we shall derive some simple approximations for r. The coefficients appearing in these approximations allow nice probabilistic interpretations. We start with introducing some background information.

If f is a probability density of some nonnegative random variable T and $\bar{f}$ its Laplace transform, then

$$(-)^n \frac{d^n\bar{f}}{ds^n}(0) = \mathcal{E}T^n\,, \tag{2.A1}$$

where $\mathcal{E}$ denotes the expectation operator. Below we shall use the standard notation $\mu:=\mathcal{E}T$ and $\sigma^2:=varT=\mathcal{E}T^2-(\mathcal{E}T)^2$.

EXERCISE 2.A1: Check (2.A1).

The cumulants κ_n of T are defined as

$$\kappa_n:= (-)^n \frac{d^n\log\bar{f}}{ds^n}(0)\,. \tag{2.A2}$$

κ_1 and κ_2 just equal μ and σ^2 respectively. κ_3/σ^3 is often used as a measure of the skewness of a distribution, and κ_4/σ^4 as a measure of its kurtosis.

In renewal theory our main interest is $\bar{\phi}$, the Laplace transform of the birth kernel. We have already seen that $\bar{\phi}(0)$ can be interpreted as the mean number R of offspring of an individual. The normalized kernel

$$f:= R^{-1}\phi \tag{2.A3}$$

can be interpreted as the probability density of a random variable, called "the age at child bearing".

REMARK 2.A2: For a biological interpretation of "the age at child bearing" consider the following thought experiment.

Start with a cohort of (potential) mothers. Wait till all their children are born. Sample at random one child. The age of its mother at the time of its birth is "the age at child bearing".

With our new notation in mind we now take a closer look at the characteristic equation (2.2.2):

$$1 = R\bar{f}(r)\,. \tag{2.A4}$$

Taking logarithms yields

$$0 = \log R + \log\bar{f}(r) = \log R - \mu r + \tfrac{1}{2}\sigma^2 r^2 + o(r^2)\,. \tag{2.A5}$$

We already know that $r=0$ for $\log R=0$. Therefore we may try to expend r in a power series in $\log R$: $r=a_1\log R+a_2(\log R)^2+\cdots$. Inserting this into (2.A5) and equating the coefficients of the successive powers of $\log R$ to zero gives

$$r = \frac{\log R}{\mu}\,(1+\tfrac{1}{2}(\sigma/\mu)^2\,\log R)+o(\log R))\,. \tag{2.A6}$$

(2.A6) highlights the effects of the mean number of offspring, mean age at childbearing and variance of the age at child bearing on the population growth rate.

Another approximation which is more accurate when the higher order cumulants are relatively small (or equivalently when the distribution of the age at child bearing is close to normal) comes from simply neglecting the $o(r^2)$ term in (2.A5), and solving the resulting quadratic equation for r:

$$r \simeq \frac{\mu-\sqrt{\mu^2-2\sigma^2\log R}}{\sigma^2}\,. \tag{2.A7}$$

The main usefulness of (2.A6) and (2.A7) is that they may do away with the need to measure the whole kernel ϕ. Generally simple statistics like R, μ and σ^2 are much easier to obtain experimentally!

REMARK 2.A3: If we are outside the range of applicability of the approximation formulae the only thing we can do is fit some low parameter family of curves to the observed individual reproduction data. A family with an exceptionally good record in this respect are the so-called shifted gamma densities

$$\phi(a) = \begin{cases} 0 & \text{for } a\leqslant\tau \\ R\dfrac{\alpha(\alpha(a-\tau))^{k-1}e^{-\alpha(a-\tau)}}{\Gamma(k)} & \text{for } a\geqslant\tau \end{cases} \tag{2.A8}$$

EXERCISE 2.A4: Let $\phi=Rf$ with f the gamma density

$$f(a) = \frac{\alpha(\alpha a)^{k-1}e^{-\alpha a}}{\Gamma(k)} \tag{2.A9}$$

Calculate lines of equal r/μ in the σ/μ versus $\log R$ plane. Compare the results with the approximate equal r/μ lines derived from either (2.A6) or (2.A7).

EXERCISE 2.A5: Let $\phi=Rf$ with f the homogeneous (α,β) density

$$f(a) = \begin{cases} (\beta-\alpha)^{-1} & \text{for } \alpha<a<\beta \\ 0 & \text{elsewhere} \end{cases}$$

Same questions as in the previous exercise.

EXERCISE 2.A6: Let ϕ be as in (2.A8). Calculate lines of equal r/τ in the $\log R$ versus $k/\alpha\tau$ plane for various values of $k^{-\frac{1}{2}}$. Compare the results with the approximate equal r/τ lines calculated from (2.A6) or (2.A7).

EXERCISE 2.A7: Devise a procedure to estimate the parameters of (2.A8) given a sample of ages at child bearing.

* EXERCISE 2.A9: Derive explicit expressions for R, μ and σ^2 for the model from example 1.4.8, using the directly available results for moments of the absorbtion time in diffusion processes (see *e.g.* GOEL & RICHTER DYN, 1974, or EWENS, 1969, 1979). (NB. The general result looks horrible. An interesting problem is whether these expressions can

be simplified for appropriate choices of ν and σ^2. It would also be of interest to derive approximation formulae based on the assumption that σ^2 is relatively small.)

2.B. The dependence of the long run population size on the initial age distribution

In subsection 2.3.1 we argued that for $t\to\infty$

$$n(t,a) = [T_2(t)n_0](a) = e^{rt}c(n_0)\pi(a)e^{-ra} + o(e^{rt}) , \tag{2.B1}$$

where n_0 is the initial age distribution. In this appendix we shall apply ourselves to calculating c direct from the interpretation. Before doing so we first introduce some notation. $n(t)$ will denote the age distribution at t. So $n(0)=n_0$. Moreover we shall normalize the eigenfunction $\tilde{n}$ corresponding to the dominant eigenvalue by setting $\tilde{n}(0)=1$, *i.e.*

$$\tilde{n}(a) = \pi(a)e^{-ra} . \tag{2.B2}$$

(Please note that in general $\int n(a)da \neq 1$.) Further notation will be introduced as the need arises.

As T_2 is a linear semigroup c has to be a linear functional on the space of age distributions. Therefore we try writing

$$c(n) = \int_0^{a_m} \kappa\nu(a)n(a)da , \tag{2.B3}$$

with κ a constant introduced to absorb any future normalization.

* REMARK 2.B1: If, as we did in subsection 2.3.1, we choose our p-state space to be $L_1(0,a_m)$ any linear functional c can be written in the form (2.B3) with $\nu \in L_\infty(0,a_m)$. The explicit formula for ν to be described below shows that when $\mathcal{F}$ is continuous, ν is continuous as well. Therefore under this condition (2.B3) also applies to an enriched space containing in addition measures such as delta "functions".

Now consider a population started with one newborn individual. Let us call the age distribution of the clan she founds $n_{0,0}$. Eventually this clan will grow exponentially, so, with an appropriate definition of κ

$$n_{0,0}(t) \simeq \kappa\tilde{n}e^{rt} , \quad t\to\infty , \tag{2.B4}$$

where κ still has to be determined. Given (2.B4) we can calculate the asymptotic size of the clan founded by an individual aged a freshly introduced at time zero. Let $n_{\tau,a}$ denote the age distribution of the clan founded by a daughter born from her at age τ and time $\tau-a$. Then

$$n_{\tau,a}(t) \simeq \kappa\,\tilde{n}\,e^{r(t-\tau+a)} . \tag{2.B5}$$

The probability that an individual aged a ever reaches age τ is $\pi(\tau)/\pi(a)$, and the probability per time unit that she begets a daughter at that age is $\lambda(\tau)$. Therefore, if we add up all descendants of our individual aged a at time zero we get

$$n_{+,a}(t) = \int_a^{a_m} n_{\tau,a}(t)\lambda(\tau)\pi(\tau)/\pi(a)d\tau \simeq \kappa\tilde{n}\nu(a)e^{rt} , \tag{2.B6a}$$

with

$$\nu(a) := \frac{e^{ra}}{\pi(a)} \int_a^{a_m} \lambda(\tau)\pi(\tau)e^{-r\tau}d\tau . \tag{2.B6b}$$

Therefore

$$n(t) = \int_0^{a_m} n_{+,a}(t)n_0(a)\,da \simeq \kappa \int_0^{a_m} \nu(a)n_0(a)da\ \tilde{n}e^{rt} . \tag{2.B7}$$

The only thing that remains is finding an expression for κ. To that end we observe that necessarily

$$T_2(t)\tilde{n} = \tilde{n}e^{rt} = c(\tilde{n})\tilde{n}e^{rt} . \tag{2.B8}$$

Therefore we should have that $c(\tilde{n})=1$ or

$$\kappa = (\int_0^{a_m} v(a)\tilde{n}(a)da)^{-1} = (\int_0^{a_m} a\ \mathfrak{P}(a)e^{-ra}\ \lambda(a)da)^{-1}\ . \tag{2.B9}$$

The function v introduced in (2.B6b) is known in the biological literature as the *reproductive value* of an individual aged a. The reason is that for large t the size of a clan founded by an individual aged a at time zero equals $v(a)$ times the size of the clan founded by a newborn individual.

EXERCISE 2.B2: Verify the last statement.

EXERCISE 2.B3: Confirm (2.B9). Hint: Use partial integration.

EXERCISE 2.B4: Derive the coefficient preceding e^{rt} in the renewal theorem 2.2.7, or in the right hand side of Feller's renewal theorem 2.2.8, by the same sort of procedure as used to derive $c(n_0)$.

EXERCISE 2.B5: Show that V defined by

$$V(t) := c(n(t)) \tag{2.B10}$$

satisfies

$$\frac{dV}{dt} = rV \tag{2.B11}$$

Hint: Differentiate under the integral sign in the definition of $c(n)$ and use the partial differential equation. Integrate partially and use the side condition.

* EXERCISE 2.B6: Show that the calculations of the previous exercise are equivalent to checking that

$$A_2^* c = rc \tag{2.B12}$$

where A_2^* is the adjoint of A_2.

3. Extensions of the linear theory

3.0. Introduction

In the introduction to this chapter we argued for the generality of the age dependent formalism in the linear case. The only thing that matters is a particular kind of "decoupling property": all individuals should be born equal, *i.e.* be independent of their forebears. This allows for a simple bookkeeping, in particular of the births, but also of derived quantities like total population size (compare formula (1.3.2)). It generally pays to keep an open mind for potential decoupling properties of this kind as their presence may be revealed or hidden depending on the way a population problem is formulated at the start. As all modelling involves an element of simplification it is even possible to introduce or do away with such properties before any equations are written down!

REMARK 3.0.1: We find it difficult to give explicit expression to the intuition underlying the general concept of "decoupling property" as it is used by us here. Roughly speaking we use the term to refer to various exploitable kinds of (stochastic) independencies inherent in *i*-behaviour.

In this section we present three examples of population problems which can be analyzed in great detail by relying on particular decoupling properties. In dealing with these examples we shall immediately use the relevant decoupling properties as our starting point. Especially in the last example, however, we only stumbled on the appropriate decoupling properties after a not inconsiderable number of false starts.

All three examples consider the distribution of some immediately observable discrete characteristic, like the number of scars on a cell derived from the division process or the size of small colonies of cells. As presented the examples may appear a trifle specific. However, they are representative of a large class of problems of their general kind. The specific details of each problem may depend on the very details of the biology of the organism under consideration, the method of approach, however, is (fairly) general.

3.1. Scar distributions in yeast

3.1.1. Budding yeasts

A full grown cell of the budding yeast *Saccharomyces cerevisiae* reproduces by forming small buds, which, after a short period of growth while remaining attached to the mother, detach themselves and start their lives as small independent cells. At the place where the bud was formed the mother cell retains a visible scar. Therefore the reproductive history of a cell can at least partially be traced by counting the number of such scars. (There even have been observed cells with over forty scars!) Given that we have available such an easy observable the questions arise (1) how the distribution of scar numbers develops over time, and (2) whether we might exploit any knowledge about the scar distribution in a population to make inferences about some aspect or other of the dynamics of the individual reproductive process. Models for scar distributions already have a not unconsiderable history in the literature. Some references are Adams *et al* (1981), HAMADA *et al* (1982), HAMADA & NAKAMURA (1982), HJORTSO & BAILEY (1983), TULJAPURKAR (1983) and GYLLENBERG (1985,a,b).

The most obvious decoupling moment in the life history of a budding yeast is the moment a young bud starts its life as an independent individual. Other decoupling moments of a different kind may occur when a yeast cell sheds a bud. We shall start concentrating on the first decoupling moment only as this allows us to write down just one renewal equation instead of a(n infinite) system of such equations. Let $\phi(a)$ denote the (mean) probability per unit of time that a yeast cell sheds a bud a time units after it started its independent life, then the rate of production of freshly detached buds b_0 satisfies

$$b_0 = \phi * b_0 + g \,, \tag{3.1.1}$$

where g as usual denotes the rate of production of buds by cells that were already present at $t=0$. If we may assume, as we shall do from now on, that each time a bud is shed the only memory that remains of the reproductive history of a yeast cell is fully parametrized by the number of scars then ϕ can be calculated as

$$\phi = \sum_{n=0}^{\infty} \phi_n * \cdots * \phi_0 \,, \tag{3.1.2}$$

where ϕ_i denotes the, possibly defective (due to deaths), probability density of the time till the next bud is shed by a cell with i scars.

EXERCISE 3.1.1: Derive (3.1.2).

Our eventual aim is to calculate the scar distribution. Let b_i be the production rate of cells with i scars, $\mathcal{F}_i$ the survivor function of cells with i scars, *i.e.* $\mathcal{F}_i(\tau)$ is the probability that a cell with i scars is alive and moreover has not shed a new bud τ time units after it came into being as a cell with i scars, and let N_i denote the total number of scars, then for $i>0$

$$b_i = \phi_{i-1} * \cdots * \phi_0 * b_0 + g_i \,, \tag{3.1.3a}$$

and for all i,

$$N_i = \mathcal{F}_i * b_i + M_i \,, \tag{3.1.3b}$$

where g_i is the rate at which the remnants of the initial population are still producing cells with i scars, and M_i is the number of cells initially present which still have i scars. (If cells are certain either to die or to produce daughters, M_i will eventually go to zero for any fixed i, and the same applies to any biologically acceptable g_i.) Finally let N denote the total population size and p_i the relative frequency of cells with i-scars, then

$$N = \sum N_i \,, \qquad p_i = N_i / N \,. \tag{3.1.4}$$

EXERCISE 3.1.2: Derive (3.1.3).

EXERCISE 3.1.3: Derive the following infinite system of renewal equations for the b_i

$$b_i = \phi_{i-1} * b_{i-1} + g_i \,, \qquad b_0 = \sum_{j=0}^{\infty} (\phi_j * b_j + g_j) \,, \tag{3.1.5}$$

where we conventionally put $g_0=0$. Use (3.1.5) to rederive (3.1.1) with

$$g = \sum_{n=1}^{\infty} (g_n + \sum_{j=0}^{n-1} \phi_n * \cdots * \phi_{n-j} * g_{n-j}) = \sum_{n=1}^{\infty} (g_n + \sum_{j=n}^{\infty} \phi_j * \cdots * \phi_n * g_n) \,. \tag{3.1.6}$$

Hint: Either exploit (2.1.2) and (2.1.3) or use Laplace transformation and ordinary algebra.

The characteristic equation corresponding to (3.1.1) is

$$1 = \bar{\phi}(r) = \sum_{n=0}^{\infty} \prod_{j=0}^{n} \bar{\phi}_j(r) . \tag{3.1.7}$$

The renewal theorem then tells us that our yeast population will eventually grow at a rate $r>0$, where r is the unique real solution of (3.1.7), if $\bar{\phi}(0)>1$. As the two problems which we set out to solve are only of interest for growing yeast populations we shall from now on make that assumption. Setting $b_0(t)=e^{rt}(c+o(1))$, for $t\to\infty$, in (3.13) gives

$$N_i(t) = e^{rt}(c\, \bar{\mathfrak{F}}_i(r) \prod_{j=0}^{i-1} \bar{\phi}_j(r) + o(1)) , \tag{3.1.8}$$

and therefore, when $\tilde{p}_i$ denotes the asymptotic value of p_i for large t,*

$$\tilde{p}_i \propto \bar{\mathfrak{F}}_i(r) \prod_{j=0}^{i-1} \bar{\phi}_j(r) , \tag{3.1.9}$$

where the product is interpreted as 1 when $i=0$. This solves our first problem.

If we are to make any progress on the inverse problem *(viz.* how can (3.1.9) be used to extract information about individual reproductive behaviour from the observed $\tilde{p}_i$) we have to make some further simplifying assumptions. Our first assumption, well entrenched in the microbiological literature, is that yeast cells do not die from old age or accumulation of scars, so that the only source of cell losses are random accidental deaths, or random washout from a culture vessel. Let μ denote the (age independent) death rate and ψ_i the probability density of the time till the next bud is produced by a cell already having i scars when there are no deaths, then

$$\phi_i(\tau) = e^{-\mu\tau} \psi_i(\tau) , \tag{3.1.10}$$

$$\mathfrak{F}_i(\tau) = e^{-\mu\tau} \int_{\tau}^{\infty} \psi_i(\sigma) d\sigma , \tag{3.1.11}$$

and

$$\bar{\phi}_i(s) = \bar{\psi}_i(\mu+s) , \tag{3.1.12}$$

$$\bar{\mathfrak{F}}_i(s) = \frac{1-\bar{\phi}_i(s)}{\mu+s} . \tag{3.1.13}$$

EXERCISE 3.1.4: Derive (3.1.10) and (3.1.11). Hint: Use the independence of the reproductive mechanism and cell loss.

EXERCISE 3.1.5: Derive (3.1.12) and (3.1.13).

Inserting (3.1.13) in (3.1.9) and setting the sum of the $\tilde{p}_i$ equal to one now gives

$$\tilde{p}_i = (1-\bar{\phi}_j(r)) \prod_{j=0}^{i-1} \bar{\phi}_j(r) . \tag{3.1.14}$$

where again the product is interpreted as 1 when $i=0$.

Using the explicit expression (3.1.14) in combination with the characteristic equation (3.1.7) it is easy to verify that the mean number of scars

$$\sum_{i=0}^{\infty} i\, \tilde{p}_i = 1 \tag{3.1.15}$$

in accordance with elementary biological reasoning: for every cell that is produced exactly one scar is produced as well. If a different mean scar number is found then we may conclude that cell death rate depends on scar number or cell age. (The other explanations that the sample is biased or that the stable value of the mean scar number has not yet been reached are both extremely unlikely, due to the basic nature of the biological mechanism of one new cell one scar. This acts equally, locally in non well stirred vessels, in growth limited populations, and by extension also in the population from which the inoculum was derived.)

* $\propto$ means proportional to.

EXERCISE 3.1.6: Check (3.1.15) algebraically.

To get some further idea of how (3.1.14) can be used assume temporarily that $\psi_i = \psi_0$ for all i. In that case the characteristic equation reduces to

$$1 = \sum_{n=0}^{\infty} [\bar{\phi}_0(r)]^{n+1} = \frac{\bar{\phi}_0(r)}{1-\bar{\phi}_0(r)} . \tag{3.1.16}$$

Therefore $\bar{\phi}_0(r) = \frac{1}{2}$ and consequently

$$\tilde{p}_i = (\tfrac{1}{2})^{i+1} . \tag{3.1.17}$$

If (3.1.17) is found to apply in our culture we may safely conclude that there is no relation between the number of scars and the time till the appearance of the next bud. In that case the scar distribution also cannot tell us anything further about the reproductive mechanism. This last conclusion may not come as a surprise. However, the main importance of (3.15) is as a null model, against which to test alternative hypotheses.

EXERCISE 3.1.7: Derive (3.1.17) by the following direct argument. Concentrate on the p_i. As deaths occur through random removal the dynamics of the p_i is not influenced by the deaths. Therefore any statement pertaining to the p_i should be independent of μ. Choose μ so that $r=0$. Then at equilibrium the N_i are constant so we can think in terms of the $\tilde{N}_i$ instead of the $\tilde{p}_i$. Therefore the transfer into the i-th scar class exactly balances the transfer and deaths from that scar class. When the ϕ_i are equal the relative transfer and death rates at equilibrium should be equal for all scar classes. This allows us to conclude that at equilibrium $\tilde{N}_i = \alpha \tilde{N}_{i-1}$ for some α, and therefore $\tilde{p}_i = \alpha \tilde{p}_{i-1}$. α can be determined from $\Sigma i \tilde{p}_i = 1$.

The next most simple model biologically is that $\psi_i = \psi_1$ for all $i \geqslant 1$ but that ψ_0 is different. In that case the characteristic equation reduces to

$$1 = \bar{\phi}_0(r) \sum_{n=0}^{\infty} [\bar{\phi}_1(r)]^n = \frac{\bar{\phi}_0(r)}{1-\bar{\phi}_1(r)} , \tag{3.1.18}$$

and we obtain

$$p_0 = 1-\bar{\phi}_0(r) = \bar{\phi}_1(r) \qquad p_i = [\bar{\phi}_0(r)]^2 [\bar{\phi}_1(r)]^{i-1} . \tag{3.1.19}$$

Therefore, if (and only if) newborn cells differ from full grown cells in their capacity to reproduce, but from the first reproduction act onwards cells stay essentially the same reproductively, the scar distribution is geometric with a modified zero'th term, and this zero'th term is related in a very specific way to the geometric parameter. Figure 3.1.1 shows the fit of this theoretical distribution to an observed one, drawn from HAMADA *et al* (1982). It can be seen that (3.1.17) certainly is out of the question, but that the fit of (3.1.19) is exceedingly good.

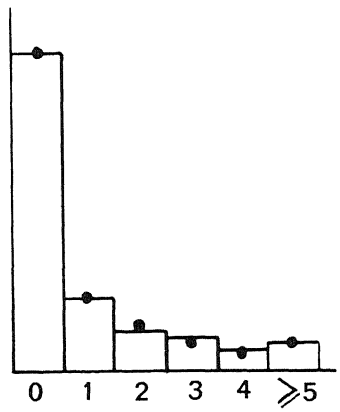

Fig. 3.1.1: Observed and theoretical •, based on (3.1.19) with $\bar{\phi}_1(r) = 0.6172$, scar number distribution in a sample of budding yeast *(Saccharomyces cerevisiae)* cells. Data from HAMADA *et al* (1982).

A final question is what kind of information about the difference between ψ_0 and ψ_1 can be obtained from

$\bar{\phi}_1(r)=1-\bar{\phi}_0(r)$ as estimated from the data. A biologically plausible hypothesis is that

$$\psi_0(\tau) = \begin{cases} 0 & \text{for } \tau<T \\ \psi_1(\tau-T) & \text{for } \tau\geqslant T \end{cases}, \tag{3.1.20}$$

where T is the time needed for the young bud to grow to an appropriate size. Then

$$\bar{\phi}_0(s) = e^{-(\mu+s)T}\bar{\phi}_1(s), \tag{3.1.21}$$

and

$$T = (\mu+r)^{-1} \log\left[\frac{p_0}{1-p_0}\right]. \tag{3.1.22}$$

As r is just the population growth rate which can be relatively easily determined, and as generally random deaths can be neglected so that μ equals the dilution rate, we may conclude that the scar number distribution just enables us to calculate the time needed for a bud to grow to reproductive size, but no more than that.

Concluding remarks

We have shown how a judicious exploitation of various types of presumed "decoupling" moments in the life of a budding yeast cell allows us to derive relatively easily an explicit expression for the asymptotic distribution of the number of bud scars. The main assumption was that budding depends only on the time since the last bud was shed and the number of scars present, and not *e.g.* directly on cell age. When we moreover made the further (decoupling) assumptions that death rate μ is constant, independent of age or scar class, and that the interbudding times are identically distributed, independent of scar number, then the stationary scar distribution was found to be geometric with a modified zeroth term. The size of this zeroth term depends in a subtle manner on the distributions of both the interbudding times and the time till the production of the first bud. If the time till the first bud only differs from the interbudding times by the addition of a fixed delay T, the size of the zeroth term p_0 could be related to $(\mu+r)T$, where r is the population growth rate, allowing T to be inferred when p_0, r and μ are known (or, conversely, to infer r when T and μ are known *etc.*). These results indicate how a relatively easily obtainable population datum like the scar number distribution can be exploited to make inferences about the behaviour of the individuals constituting that population.

In all our calculations we concentrated on stationary distributions only, as this is the kind of datum that is most easily obtainable experimentally. Of course more information can be gleaned, at least in principle, from transient behaviour (see *e.g.* BERAN, STREIBLOVA & LIEBLOVA (1969), HJORTSO & BAILEY (1983) and TULJAPURKAR (1983)).

Finally we refer to GYLLENBERG (1985) for an analysis of more detailed models relating the scar distribution to cell sizes as the basic variable underlying the budding process in the manner of I.4 and II.

EXERCISE 3.1.8: Argue that under the assumptions made in this subsection the state of an individual can be parametrized by its scar number i and the time τ since it last shed a bud or was born. Write down the partial differential equations for the p-state (see chapter III for the methodology). Assume initially that the death rate μ may depend on i as well as on τ.

EXERCISE 3.1.9: Find an equation for the dominant eigenvalue and an expression for the corresponding stable distribution for the differential generator from exercise 3.1.8.

EXERCISE 3.1.10 Derive (3.1.9) from the result of exercise 3.1.9 with appropriate definitions of $\bar{\phi}$ and $\bar{\mathcal{F}}$.

EXERCISE 3.1.11: Now assume that $\mu_i(\tau)=\mu$ and derive (3.1.14) to (3.1.19)

3.1.2. Fission yeasts

Fission yeasts, *Schizosaccharomyces pombe,* differ from budding yeasts in that they reproduce by cells dividing into two approximately equal parts. In the process two scars are formed, one on each sister cell. Moreover the scars that were already present on the mother are distributed over the daughter cells. The precise nature of the distribution rules is unknown. In fact the biological question is how the observed scar distribution can be used to make inferences about these rules and therefore by extension about the underlying process of cell growth and division (see *e.g.* CALLEJA *et al.* (1980) and HAMADA (1982)). Our main reason for including precisely this example, however, is technical: it shows how yet another type of biologically plausible decoupling property can be exploited to obtain more or less explicit mathematical results.

For a change we shall start this time from the partial differential equations for the p-state, assuming that the i-state is fully parametrized by the scar number i and cell age a:

$$\frac{\partial n_i(t,a)}{\partial t} = -\frac{\partial n_i(t,a)}{\partial a} - (d_i(a) + \mu_i(a))n_i(t,a)\,, \tag{3.1.23}$$

$$n_i(t,0) = 2 \sum_{j=max(1,i-1)}^{\infty} \pi_{ij} \int_0^{\infty} d_j(\alpha)n_j(t,\alpha)d\alpha\,,$$

where π_{ij} denotes the probability that a randomly chosen daughter of a cell with j scars has itself i scars, and $d_i(a)$ and $\mu_i(a)$ are the probabilities per unit of time that a cell with i scars aged a divides resp. dies.

The scar distribution can be found from

$$N_i(t) = \int_0^{\infty} n_i(t,a)da, \qquad N(t) = \sum_{i=1}^{\infty} N_i(t)\,, \qquad p_i(t) = N_i(t) \;/\; N(t)\,. \tag{3.1.24}$$

From the interpretation it follows immediately that

$$\pi_{ij} = \pi_{j+2-i,j}, \quad \text{for } i = 1, \cdots, j+1\,, \tag{3.1.25a}$$

$$\sum_{i=1}^{j+1} \pi_{ij} = 1\,. \tag{3.1.25b}$$

The values of the π_{ij} will be determined in some rough fashion by the processes of cell growth and division. In particular scar redistribution will be much more uneven when cell growth is unipolar than when it is bipolar *(vide* CALLEJA *et al,* 1980). Provided all μ_i are zero the π_{ij} together form the transition matrix of the Markov chain that results from following one cell line down through the generations. For later reference we shall call this chain M.

To determine the behaviour of the n_i for large time we turn, as usual, to the characteristic equation derived from (3.1.23):

$$\frac{d\tilde{n}_i}{da} = -(d_i(a) + \mu_i(a) + r)\,\tilde{n}_i(a) \tag{3.1.26a}$$

$$\tilde{n}_i(0) = 2 \sum_{j=\max(1,i-1)}^{\infty} \pi_{ij} \int_0^{\infty} d_j(\alpha)\tilde{n}_j(\alpha)d\alpha \tag{3.1.26b}$$

which on setting

$$\mathcal{P}_i(a) := e^{-\int_0^a (d_i(\alpha)+\mu_i(\alpha))d\alpha}\,, \qquad \phi_i(a) := d_i(a)\,\mathcal{P}_i(a) \tag{3.1.27}$$

(the similarity of the notation to that in the previous subsection is not incidental!) reduces to

$$\tilde{n}_i(0) = 2 \sum_{j=\max(1,i-1)}^{\infty} \pi_{ij}\,\bar{\phi}_j(r)\tilde{n}_j(0) \tag{3.1.28}$$

EXERCISE 3.1.12: Show that with the definitions (3.1.27) equation (3.1.28) is equivalent to (3.1.26).

It can be shown that under fairly general conditions on the π_{ij}, in particular that any state in the Markov chain M can be reached from every other state, there exists a unique real eigenvalue r with corresponding nonnegative eigenvector. Below we shall indicate how this eigenvector can be calculated explicitly if we make some special assumptions. Comparing the observed and predicted scar distributions then provides a test of our assumptions. If some more complicated assumptions are made one has to resort to numerical methods (see *e.g.* HAMADA, 1982). But even in that case the analytical reasoning makes it possible to concentrate on a relatively simple set of equations instead of having to solve the full partial differential equations (3.1.26).

We begin by making the biologically plausible assumption (compare the previous subsection) that $\mu_i(a)=\mu$, independent of i or a. In that case the biological mechanism "one additional cell → two more scars", leads us to infer that asymptotically the mean number of scars on a cell should be two:

$$\sum_{i=1}^{\infty} i\tilde{p}_i = 2 \tag{3.1.29}$$

This provides a test whether death rate is independent of scar number. (The fact that (3.1.29) is satisfied does not rule out that cell death depends on cell age as in this example cell age is not related to scar number!) An algebraic

verification of (3.1.29) is provided by the exercises.

EXERCISE 3.1.13: Check (3.1.29). Hint: As a start show that

$$\tilde{N}_i = \int_0^\infty \tilde{n}_i(a)da = \frac{1}{r+\mu}\,\tilde{n}_i(0)\,(1-\bar{\phi}_i(r))\,,$$

then use (3.1.25) to show that

$$\sum_{i=1}^{j+1} i\,\pi_{ij} = \frac{j+2}{2}\,,$$

and use (3.1.28) to show that

$$\sum_{i=1}^{\infty} \tilde{n}_i(0) = 2\sum_{j=1}^{\infty} \bar{\phi}_j(r)\,\tilde{n}_j(0)$$

and

$$\sum_{i=1}^{\infty} i\,\tilde{n}_i(0) = \sum_{j=1}^{\infty} j\,\bar{\phi}_j\,(r)\,\tilde{n}_j(0) + \sum_{i=1}^{\infty} \tilde{n}_i(0)\,.$$

Use these identities to prove that

$$\tilde{N} = \sum_{i=0}^{\infty} \tilde{N}_i = \frac{1}{2(r+\mu)} \sum_{i=1}^{\infty} \tilde{n}_i(0)$$

and

$$\sum_{i=1}^{\infty} i\,\tilde{N}_i = \frac{1}{r+\mu} \sum_{i=1}^{\infty} \tilde{n}_i(0)\,.$$

Given the assumption of constant μ there are two types of scar redistribution rules for which we can get an expression for the p_i without making any further assumptions about the ϕ_j. The first one is

$$\pi_{1j} = \pi_{j+1,j} = \tfrac{1}{2}, \qquad \pi_{ij} = 0 \text{ for } i \neq 1,\ j+1\,, \tag{3.1.30}$$

or, in other words, one sister always ends up with one scar, the other gets all the old scars plus one new one. In essence this is the budding yeast model all over again, except that the number of scars is increased by one.

The second special case is that of equal sharing of the scar load:

$$j \text{ odd: } \pi_{ij} = \begin{cases} \frac{1}{2} & \text{for } i = (j+1)/2 \text{ and } i = (j+1)/2 + 1\,, \\ 0 & \text{for all other } i \end{cases} \tag{3.1.31}$$

$$j \text{ even: } \pi_{ij} = \begin{cases} 1 & \text{for } i = j/2 \\ 0 & \text{for all other } i \end{cases}$$

Clearly only a scar load of 2 reproduces itself in its daughters, and all cells with other scar loads give birth to daughters with scar loads nearer to 2. And indeed, from (3.1.28) we can deduce that

$$\tilde{n}_i(0) = \begin{cases} c & \text{(a free parameter) for } i = 2 \\ 0 & \text{for all other } i \end{cases} \tag{3.1.32}$$

and

$$1 = 2\,\tilde{\phi}_2\,(r)\,, \tag{3.1.33}$$

i.e. the stable distribution consists of 2-scar cells only, and the characteristic equation equals that of a population consisting wholy of selfreproducing 2-scar cells.

EXERCISE 3.1.14: Derive (3.1.32) and (3.1.33) from (3.1.28).

Returning to general π_{ij} we make as our final simplification the biologically rather plausible decoupling assumption that time to cell division is independent of scar number so that the two processes of reproduction and scar

redistribution can be considered separately. Summing over i in (3.1.28) then gives

$$1 = 2\,\bar{\phi}\,(r)\,, \tag{3.1.34}$$

the usual characteristic equation of the pure cell kinetics model. Substituting this in (3.1.28) again and using $\tilde{p}_i \propto \tilde{n}_i(0)$, results in

$$\tilde{p}_i = \sum_{j=\max(1,i-1)}^{\infty} \pi_{ij}\,\tilde{p}_j\,, \tag{3.1.35}$$

which is just the equation for the stationary distribution of the Markov chain M. Naturally in this case scar distribution does not contain any information about the division times. Using specific assumptions for the π_{ij} we can now calculate the $\tilde{p}_i$, either numerically or in some simple cases analytically. Translating hypotheses about the cell growth and division morphology into specific numerical values of the π_{ij} and subsequent calculation of the $\tilde{p}_i$ then makes possible a relatively easy empirical test of those assumptions. See CALLEJA *et al* (1980), and HAMADA (1982) for some further discussion.

EXERCISE 3.1.15: Let $\pi_{ij}=(j+1)^{-1}$ (all scar redistributions equally probable). Calculate the $\tilde{p}_i$ from (3.1.35) together with the conservation condition $\Sigma\tilde{p}_i=1$.
Hint: use (3.1.35) to derive the difference equation $\tilde{p}_i=\tilde{p}_{i-1}-(i-2)^{-1}\,\tilde{p}_{i-2}$) and use this to show that the $\tilde{p}_i=e^{-1}/(i-1)!$, *i.e.* scar number minus one is Poisson distributed with mean 1.

EXERCISE 3.1.16: Let $\pi_{ij}=q/2$, independent of j for $j\geqslant 2$ (check that necessarily $\pi_{11}=\pi_{21}=\frac{1}{2}$). In words: the probability that all the old scars end up on one of the two daughters equals q, independent of the number of old scars. Show that $\tilde{p}_1=\dfrac{q(1-\bar{\phi}_1(r))}{1+(q-1)\bar{\phi}_1(r)}$, and, therefore if $\phi_i=\phi$ independent of i, $\tilde{p}_1=q/(1+q)$.
Hint: exploit (3.1.28) and some of the identities from exercise 3.1.13.

EXERCISE 3.1.17: Write down an (infinite) system of renewal equations for $b_i(t)=n_i(o,t)$, either direct from the interpretation or by integrating (3.1.23) along the characteristics. Use the reasoning from the start of section 2.2 to arrive at (3.1.28) again.

3.2. *Colony size in the diatom* Asterionella

This example has already figured heavily in chapter III. However, except for some of the exercises, the development in this chapter has been kept completely independent, even to the extent of reintroducing the basic biology.

The planctonic diatom *Asterionella,* like many other such diatoms, occurs in small colonies. In the case of the, starshaped, colonies of *Asterionella* the main colony sizes are one, two, four, eight and sixteen. These colony sizes presumably derive from synchronous division of the cells in a colony, doubling colony size as newly formed sisters remain attached by inanimate bonds, and a breaking of the colonies into equal parts due to a weakening of those same bonds over time. In a colony of, say, size eight, one therefore has three types of bonds: four bonds linking sisters, two bonds that now link nieces, but originally linked their mothers, and one bond linking what have now become grand-nieces. (If you don't get the picture you might glance ahead to fig. 3.2.4). Below we shall refer to these bonds as respectively the child, parent and grandparent bonds. The age of any child bond equals colony cell age. If our colony of size eight does not double first, thereby transforming children into parents etc., it will in all probability fall apart into two colonies of size four each, due to a breaking of the grandparent bond.

In fast growing populations colony size generally is larger than in slow growing or declining populations. This is easily explained if under better circumstances interdivision times are shorter, but bond weakening always proceeds at the same pace. The nice thing about this is that it allows us to infer some population *growth* characteristic from a *single,* nonrepeated, sample from that population. The first mathematical problem then is to find out how exactly division and breakage rules are related to colony size distribution.

In chapter III we formulated three possible models for the division and breakage mechanisms. All these models had in common that 1) cell division rate only depended on cell age, and 2) that the probability per unit of time that a particular bond breaks depends only on the age of that bond. In the first model we assumed that a bond always breaks after a fixed time A. The second and third model allow for some variability in breakage time. In the second model bonds were assumed to break in complete dependence: if a particular bond breaks all its sisters break as well, and so do its forebears, thus keeping colony size strictly in the main sequence of 2^n, $n=0,1,2,...$ The third model

specifically was designed to allow for strays. In it the opposite assumption was made that bonds are totally independent in their breaking behaviour. (Note that the first model is a special case of both the second and third models.) Studying more complicated models, allowing a partial dependence between breakage times in a colony, makes no sense at this stage as it will never be possible to distinguish between them on such meager data as colony size distributions alone. The situation might change though when more detailed culture data become available.

All the results in this subsection are our own and have not been published before. The initial impetus came from ideas and experiments of Marcel Donze (see KOOYMAN, 1976, and WESSEL, 1984). The work is certainly not finished: the doubly starred exercises represent routes we yet want to explore. However, we considered the story so appropriate to these lecture notes that we decided to incorporate it as just a snapshot of some of our current work.

Our assumptions about cell division imply that the total birth rate of cells satisfies the usual renewal equation. We shall assume that cells do not die. Then

$$b(t) = 2 \int_0^t g(a)\, b(t-a)da + u(t)\,, \tag{3.2.1}$$

where g is the probability density of the time till division. (The reason for the somewhat aberrant notation is that in this way we are better equipped for the probabilistic arguments below.) And for large t

$$b(t) = ce^{rt} + o(e^{rt})\,, \tag{3.2.2}$$

with r the unique real root of

$$1 = 2 \int_0^\infty g(a)e^{-ra}da\,. \tag{3.2.3}$$

EXERCISE 3.2.1: How is g related to the division rate d as introduced in chapter III formulae (5.4.2), (C.6) and (C.7)? ** How is u related to the quantities occurring in these equations?

We start considering the model in which bonds break at exactly age A. To calculate $N_k(t)$, the number of colonies with size 2^k, $k>0$, at time t, we observe that the oldest bond in the colony has age $a_k<A$, k the "generation-index", and that its parent was born at some time τ ago with $\tau>A$ (see fig. 3.2.1).

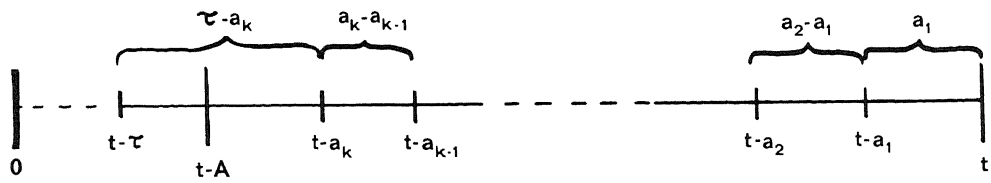

Fig. 3.2.1: Moments at which the bonds in a colony of size k and their last broken forebear were formed, and corresponding interdivision times.

The probability that exactly $k-1$ more divisions occurred between $t-a_k$ and t, so that colony size is 2^k, equals the probability that at least $k-1$ divisions occurred minus the probability that at least k divisions occurred:

$$P\ \{\text{exactly } k-1 \text{ divisions in } (t-a_k,\ t)\} = G_{k-1}(a_k) - G_k(a_k)\,,$$

with

$$G_0(a) = 1, \quad \text{and} \quad G_h(a) = \int_0^a g^{h*}(\alpha)d\alpha\,. \tag{3.2.4a}$$

Therefore, integrating over all possible combinations of birth time of the oldest bond and its parent, and observing that on breaking of the parent two colonies were formed, for $k>0$ and $t>A$,

$$N_k(t) = \int_0^A 2 \int_A^t b(t-\tau)g(\tau-a_k)d\tau\,[G_{k-1}(a_k) - G_k(a_k)]da_k + R_k(t)\,, \tag{3.2.4b}$$

where R_k refers to all colonies of size 2^k in which the parent of the presently oldest bond was born before $t=0$. To calculate N_0, the number of single cells, we observe that these cells all have age $a_0 > A$. Therefore, for $t > A$,

$$N_0(t) = 2 \int_A^t b(t-a_0)\, \mathcal{G}(a_0) da_0 + R_0(t)\,, \tag{3.2.4c}$$

where $\mathcal{G}(a_0) = 1 - G_1(a_0)$ is the probability that no further division has occurred after $t - a_0$.

EXERCISE 3.2.2: Argue that when g is not defective $R_k(t) \to 0$ for $t \to \infty$. Hint: observe that $R_k \leq \sum_m R_m$ and interpret this sum.

EXERCISE 3.2.3: Observe that for $t \geq A$ (3.2.4) can be written as

$$N_k = \psi_k * b + R_k\,. \tag{3.2.5}$$

Give expressions for the kernels ψ_k.

For large t we find on setting $p_k = N_k / \sum_m N_m$ and $\tilde{p}_k = \lim_{t\to\infty} p_k$

$$\tilde{p}_0 = c \int_A^\infty e^{-r\tau}\, \mathcal{G}(\tau) d\tau\,, \tag{3.2.6a}$$

$$\tilde{p}_k = c \int_A^\infty e^{-r\tau} \int_0^A g(\tau - a)\, [G_{k-1}(a) - G_k(a)] da\, d\tau \quad \text{for } k > 0\,,$$

with

$$c^{-1} = e^{-rA} \int_0^\infty e^{-r\tau}\, \mathcal{G}(\tau) d\tau\,. \tag{3.2.6b}$$

EXERCISE 3.2.4: Check (3.2.6). Hint: First observe that $p_k \propto N_k$, then take limits in (3.2.4) after some convenient renormalization. Finally check that for c given by (3.2.6b) $\Sigma \tilde{p}_k = 1$.

For various special choices of g, like the gamma distributions from exercise 2.A4, it is possible to evaluate (3.2.6) explicitly. We shall leave this as an exercise, but for one exception. When $g(a) = \delta(a - D)$, we find relatively easily for $\alpha := A / D$, *i.e.* α^{-1} equals the population doubling time expressed in bond life times,

$$\left.\begin{aligned} \tilde{p}_h &= (\tfrac{1}{2})^{h-\alpha-1} - 1 \\ \tilde{p}_{h-1} &= 2 - (\tfrac{1}{2})^{h-\alpha-1} \end{aligned}\right\} \text{for } \alpha < h < \alpha + 1\,,$$
$$\tilde{p}_k = 0 \qquad \text{for } k \neq h,\ h-1\,. \tag{3.2.7}$$

So given the $\tilde{p}_k$ we can immediately calculate α_1. However, some care is needed in interpreting (3.2.7). For g a delta function at D any births at t give rise to exactly twice that many births at $t + D$ without any "smearing" out. The renewal theorem as formulated in section 2 does not apply, and there is no convergence to a smooth exponential behaviour (also compare II.11). Therefore (3.2.7) cannot be interpreted as the long run values of the p_k for a proces with this division time distribution. Instead it should be interpreted as the limit for the coefficient of variation goir to zero of the long run limiting values of the p_k for some process in which there is still some variation in the interdiv sion times. In this sense (3.2.7) provides a useful approximation against which to check empirical data. The outcom from such a comparison not unexpectedly is that there clearly is a great deal of variation in either the interdivision oı breakage times, as even in exponentially growing cultures the colony sizes certainly are distributed over a much larger number of classes than two only. So if we wish to fit real data we shall have to use one of the more elaborate families of distributions discussed in the exercises.

EXERCISE 3.2.5: Check (3.2.7). How is α related to the population growth rate r?
Hint: $g^{k*}(a) = \delta(a - kD)$. See further figure 3.2.2.

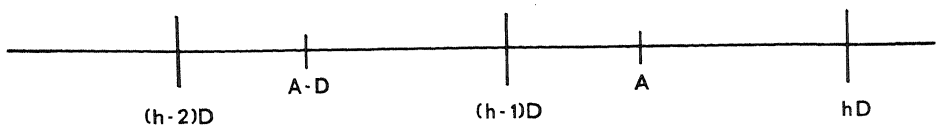

Fig. 3.2.2: The relative timing of breakage and division if the division time distribution is concentrated at D and the breakage time distribution at A.

** EXERCISE 3.2.6: Calculate explicit expressions for the $\tilde{p}_k$ assuming g to be a gamma density as in exercise 2.A4 with k an integer.
Hint: If $g_{k,\alpha}$ denotes the gamma density with parameters k and α as in exercise 2.A4, then $g_{k,\alpha}^{h*} = g_{hk,\alpha}$, and if $\mathcal{G}_{k,\alpha}$ denotes the corresponding survivor function, then for integer k

$$\mathcal{G}_{k,\alpha}(a) = \sum_{i=0}^{k-1} e^{-\alpha a} (\alpha a)^i / i ! \tag{3.2.8}$$

** EXERCISE 3.2.7: Calculate explicit expressions for the $\tilde{p}_k$ on the assumption that g is a shifted gamma density, *i.e.* $g(a)=0$ for $a<\Delta$ and $g(a)=g_{k,\alpha}(a-\Delta)$ for $a>\Delta$ (notation as in the previous exercise), with k an integer.

When bonds do not break at an exactly fixed age we have some choice in making assumptions about the dependence structure of the breakage times of the different bonds in a colony. If we wish our colony sizes to remain in the main sequence only (2^k, $k=0,1,...$) we could *e.g.* assume that only the oldest bond can break. However, apart from the distinctly unbiological flavour of this assumption, it also makes us loose all nice decoupling properties, like the one we exploited above to such good avail. An assumption which leaves that decoupling property intact is the following: Assume that there exists a monotonically increasing function b of bond age, such that the hazard rate that all bonds of age a and older break equals $b(a)$. Under this assumption the survivor function of any specific individual bond equals

$$\mathcal{R}(a) = \exp[-\int_0^a b(\alpha) d\alpha] , \tag{3.2.9a}$$

independent of the ages of the other bonds it is associated with, and the hazard-rate of the i-th generation bonds in a colony breaking but not any younger ones, equals $b(a_i)-b(a_{i-1})$. Biologically such an assumption could be defended from the viewpoint that bonds weaken over time in a deterministic fashion and that bonds break due to random shocks in which all sufficiently weak bonds in a colony break together. This model we shall now develop in more detail.

To calculate the N_k we again look at the last broken forebear of the bonds present now. Let this forebear be born τ time units ago, and let the presently oldest bond have age a. This combination of events can come about in two ways (see fig. 3.2.3): Either the last forebear broke when the presently oldest bond was not born yet, which happens with probability

$$F(\tau-a) := 1-\mathcal{R}(\tau-a) \tag{3.2.9b}$$

or it broke at some as yet unspecified time between α and $\alpha+d\alpha$ after the presently oldest bond was born, the presently oldest bond remaining intact at that time, which combination of events happens with probability $[b(\alpha+\tau-a)-b(\alpha)]\,\mathcal{R}(\alpha+\tau-a)d\alpha$. The probability that the presently oldest bond stayed intact till now in these two cases is respectively $\mathcal{R}(a)$ and $\mathcal{R}(a)/\mathcal{R}(\alpha)$. Combining the various possibilities gives

$$\begin{aligned} H(\tau,a) &:= P\{ \text{ a bond born } \tau \text{ ago and having a child } a \text{ ago has broken itself but its child is still intact } \} \\ &= F(\tau-a)\,\mathcal{R}(a) + \int_0^a [b(\alpha+\tau-a)-b(\alpha)]\,\mathcal{R}(\alpha+\tau-a)\,\frac{\mathcal{R}(a)}{\mathcal{R}(\alpha)}\,d\alpha . \end{aligned} \tag{3.2.9c}$$

Using (3.2.9c) we may proceed just as in the previous case. The result is

$$\tilde{p}_k \propto \int_0^\infty e^{-r\tau} \int_0^\tau g(\tau-a)\, H(\tau,a)\, \{G_{k-1}(a)-G_k(a)\}\, da\, d\tau \tag{3.2.9d}$$

$$\tilde{p}_0 \propto \int_0^\infty e^{-r\tau} F(\tau)\, \mathcal{G}(\tau)\, d\tau .$$

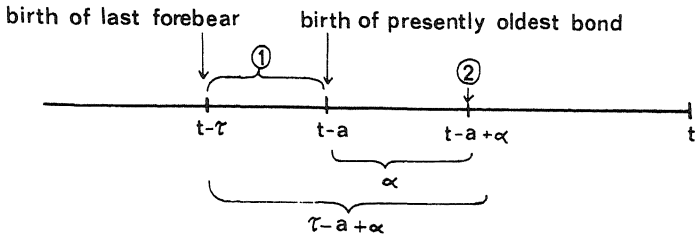

Fig. 3.2.3: Possible times at which the last forebear of the presently oldest bond could have been broken

EXERCISE 3.2.8: Check that when $\mathcal{F}(a)=1$ for $a<A$ and $\mathcal{F}(a)=0$ for $a>A$, $H(\tau,a)=1$ for $a<A<\tau$ and $H(\tau,a)=0$ elsewhere, so that for a degenerate breakage time distribution (3.2.9) indeed reduces to (3.2.6). Hint: Use delta-functions (and $0^2/0=0$).

** EXERCISE 3.2.9: Could you dream up a, non degenerate, family of distributions for which (3.2.9) can be evaluated explicitly?

In the final model to be discussed here we assume that all bonds break in complete independence. This necessarily upsets the decoupling conditions used in the previous two models. If we want to make any progress along the lines of this chapter we therefore have to introduce some other decoupling condition artificially, by restricting the supports of the probability densities g and f of respectively the time between divisions and bond lifetime. As we already know that colony sizes larger than 16 are never seen in practice, we shall assume that bonds can become at most great-grandmothers (4 generations in one colony), *i.e.* we assume that there exist a $D_{\min}$ and an $A_{\max}$, $A_{\max}<4D_{\min}$, such that

$$\left.\begin{array}{l} g(a)=0 \\ \mathcal{G}(a)=1 \\ G(a)=0 \end{array}\right\} \text{for } a<D_{\min}, \qquad \left.\begin{array}{l} f(a)=0 \\ \mathcal{F}(a)=0 \\ F(a)=1 \end{array}\right\} \text{for } a>A_{\max}, \tag{3.2.10}$$

where $\mathcal{G}$ and $\mathcal{F}$ and G and F are respectively the survivor functions and distribution functions corresponding to g and f, defined in the usual manner.

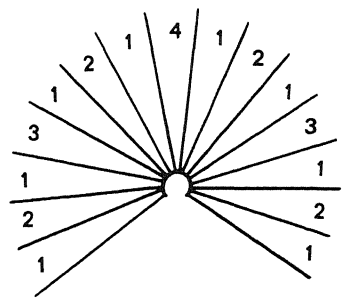

Fig. 3.2.4: (Hypothetical) colony of size 16 with the relative ages of the bonds (central) joining the radiating cells, 1 indicating the youngest bonds and 4 indicating the oldest one.

If bonds break independently colonies of all sizes may occur. Therefore we have to change our notation. M_i will denote the number of colonies of size i and q_i their relative frequency. (So $N_k=M_{2^k}$ and $p_k=q_{2^k}$). Due to the simplifying assumption (3.2.10) we can consider any colony as fragments of some hypothetical colony of size 16 (see figure 3.2.4). If the ages of the bonds in that colony are a_1 to a_4, where the numbering is as in fig. 3.2.4, then it is a simple combinatorial exercise to calculate the various ways in which a colony of size k can be obtained as a breakage product

and their corresponding probabilities. Integrating over a_1 to a_4 then finally gives:

$$\tilde{q}_i \propto \int_0^{\infty} \int_0^{a_4} \int_0^{a_3} \int_0^{a_2} e^{-ra_4} g(a_4-a_3)g(a_3-a_2)g(a_2-a_1)\mathcal{B}(a_1)C_i^4(a_1,..,a_4)da_1..da_4 \tag{3.2.11a}$$

with

$$\begin{aligned} C_{16}^4 &= \mathcal{R}(a_4)\,\mathcal{F}^2(a_3)\,\mathcal{F}^4(a_2)\,\mathcal{F}^8(a_1) \\ &\cdot \\ &\cdot \\ &\cdot \\ C_2^4 &= \mathcal{R}(a_1)\,F(a_2)[2 + 4\,F(a_3) + 2\,F(a_4)] + F^2(a_1)[4\,\mathcal{R}(a_2) + 2\mathcal{R}(a_3) + \mathcal{R}(a_4)] \\ C_1^4 &= F(a_1)[2 + 8F(a_2) + 4F(a_3) + 2F(a_4)] \end{aligned} \tag{3.2.11b}$$

EXERCISE 3.2.10: Calculate C_3^4 to C_{15}^4 by direct counting of the possibilities.

EXERCISE 3.2.11: Calculate the $\tilde{q}_i$ for g a delta function at D.

** EXERCISE 3.2.12: Let C_1^k to $C_{2^k}^k$ denote the coefficients analogous to C_1^4 to C_{16}^4 but now for the breakage products of a colony of size 2^k. Find a way to express the C_i^k, $i=1,..,2^k$ in the C_j^{k-1}, $j=1,..,2^{k-1}$. Use this recurrence relation to rederive the result of the previous exercise.
REMARK 3.2.13: Using the result of this exercise and computer methods it should be possible in principle to choose k very large, and in this way to relax the restricting conditions on which (3.2.11) was based.

** EXERCISE 3.2.14: In the three models discussed thus far we assumed that all individuals in a colony divide simultaneously. This is in accord with the observation that in a sample always either no cells in a colony are in the process of dividing or all of them are. However, models in which cells in a colony are less than fully dependent are thinkable as well. Again as an extreme one might assume that cells divide totally independently. (If the interdivision time distributions have but a small coefficient of variation and colony size is small, then cells in one colony would still tend to divide close to each other due to the trivial "synchronization" of their single common ancestor.) Try to divise models of this kind.

EXERCISE 3.2.15: How would the results obtained above be modified when a constant nonzero death rate is assumed. (Hint: Use the interpretation, there is almost no need to do any calculations!)

The three models described in this subsection represent the stage we have reached thus far with the *Asterionella* problem (or even more than that if you take account of the doubly starred exercises). The final stage of all this work should be the comparison of the calculated distributions for the various models with real, laboratory and field, data. After a model is found to fit sufficiently well the relationship of the estimated model parameters to the population growth rate should be worked out, *e.g.* by using an expansion procedure like that described in 2.A. We hope soon to find time to try our hand at all this.

You will no doubt have noticed that we did not refer even once to the differential equations for the *Asterionella* models derived in chapter III. The techniques used in this chapter simply are more powerful in this particular case. Yet it was precisely our investigation of those differential equations which gave us a sufficient grasp of the problem to come up with the solutions presented here. Moreover the fact that we could show (3.2.6) to match with the eigenfunction of (III.C.6) for the eigenvalue r gave us the confidence to proceed with the present approach.

* EXERCISE 3.2.16: Derive (3.2.6) from (III.C.6). Hint: First derive a formula for the stationary p-state distribution $\tilde{n}$ using the kind of reasoning which was used to derive (3.2.6). Check that the result satisfies $A\tilde{n}=r\tilde{n}$ where A is the differential generator from (III.C.6). Finally calculate the $\tilde{p}_i$ from the $\tilde{n}_i$.

** EXERCISE 3.2.17: Derive (3.2.9) from (III.C.7).

*** EXERCISE 3.2.18: Derive (3.2.11) from (III.5.4.2).

4. Some nonlinear extensions of the linear theory

In this section and the next one we shall consider various examples of nonlinear results which are directly based on the linear theory from section 2. The results in the present section discuss how in some cases p-behaviour far from the trivial equilibrium (the equilibrium at which $b=N=0$) can be related to the behaviour of the linear problem derived by linearizing near that equilibrium. In the next section we deal with conditions which allow the reduction of the full p-equations of age dependent population dynamics to a differential equation on a finite dimensional space. A general survey of techniques for tackling nonlinear population problems is given in chapter VI. A good mathematical discussion of the general theory of nonlinear age dependent population dynamics can be found in WEBB (1985a). HOPPENSTEADT (1975) and especially NISBET & GURNEY (1982) provide some interesting biological perspectives.

In subsection 4.1 we discuss what might well be the oldest population dynamical model formulated as a nonlinear Volterra integral equation, the Kermack-McKendrick equation for the general epidemic. The main reason for treating this example, however, is that it allows us to introduce what might be considered the natural extension of the classic renewal theorem to the nonlinear realm. In subsection 4.2 we show for the *Daphnia* example from chapters I and III how the solution of the linear renewal equation may be used to construct an estimate for the solution of a nonlinear problem, and thereby to prove global stability of the trivial equilibrium.

4.1. Kermack's and McKendrick's (1927) general epidemic and the nonlinear renewal theorem

In (1927) KERMACK and MCKENDRICK introduced a general epidemic model, the assumptions of which have been set forth in example 1.3.3. For this model they proved the famous so-called threshold theorem, and also showed how to calculate the initial rate of increase of the epidemic from a linear renewal equation obtained by linearizing their general equation around the uninfected state. Oddly enough the general Kermack & McKendrick model was largely neglected by later applied mathematicians (see remark 1.3.4). Only in the 1970's KERMACK's and MCKENDRICK's work was again brought to the attention in the work of REDDINGIUS (1971) and WALTMAN (1974). As an answer to some of the questions raised by REDDINGIUS, METZ (1978) conjectured the "nonlinear renewal theorem" to be discussed below. The first part of this theorem was subsequently proved by DIEKMANN (1977; see also DIEKMANN & KAPER, 1978, and DIEKMANN & VAN GILS, 1984), the second part by GRIPENBERG (1983c) (but with a wrong reference as to the origin of the confecture!).

The answers to the exercises in this subsection can all be found in METZ (1978) or REDDINGIUS (1971).

Let s denote the number of susceptibles, $\dot{s}$ its time derivative, $\phi(\tau)$ the mean infectivity of an individual first infected τ time units ago (ϕ is called the infectivity kernel), and $g(t)$ the infectivity at time t due to outside sources and/or individuals that were already ill before $t=0$, then Kermack's and McKendrick's equation is (see example 1.3.3).

$$\dot{s}(t)/s(t) = \int_0^t \dot{s}(t-\tau)\phi(\tau)d\tau - g(t)\,. \tag{4.1.1}$$

In words: The relative rate of change of the susceptible number equals minus the total infectivity, which in turn equals the cumulation of all the infectivities produced by individuals infected after $t=0$ (NB $\dot{s}$ is negative) plus the infectivity contributed by individuals already ill at $t=0$ and/or outside sources. Note that no new susceptibles are produced either through births or through previously infected individuals returning to the susceptible population.

EXERCISE 4.1.1: Assume that an infected organism stays ill for some random time with survivor function $\mathfrak{F}$ after which it dies or recovers and becomes immune forever, and that an ill organism which contracted its illness τ time units ago has infectivity $a(\tau)$. Assume moreover that the susceptible population does not change due to other causes than individuals succumbing to disease. Write down joint differential equations for the "age" distribution of ill organisms (where "age" refers to the illness) and the susceptible number. Derive an equation of the form (4.1.1) by solving the partial differential equation for the age distribution by integration along characteristics on the supposition that $\dot{s}$ is known (Compare example 1.4.1 and in particular exercise 1.4.2).

EXERCISE 4.1.2: Assume that an individual illness follows a k-state, continuous time Markov chain with killing, where "killing" means either to recover and be immune forever or to die. Let the differential generator of this chain be B, let a newly infected organism enter the states of illness according to a probability vector C, and let the infectivity of the various states of illness be collected in a vector A. Write down joint differential equations for s and the distribution N of ill organisms. By solving the (linear) differential equation for N as if $\dot{s}$ were known and substituting the result in the equation for s derive an equation of the form (4.1.1). (Compare example 1.4.3 and in particular exercise

1.4.4.).

For the purpose of our further calculations it is easier to transform (4.1.1) into two more manegeable forms. Integrating once we find

$$\ln(s(t)/s_0) = \int_0^t s(t-\tau)\phi(\tau)d\tau - s_0 \int_0^t \phi(\tau)d\tau - h(t) \tag{4.1.2}$$

with

$$h(t) := \int_0^t g(\tau)d\tau \tag{4.1.3}$$

A rescaling

$$p := (s_0 - s)/s_0, \qquad \gamma := \int_0^\infty \phi(\tau)d\tau, \qquad \psi := \gamma^{-1}\phi \tag{4.1.4}$$

where p is the fraction of victims and γ is the total infective strength of the disease, then gives

$$-\ln(1-p(t)) = \gamma s_0 \int_0^t p(t-\tau)\psi(\tau)d\tau + h(t)\,. \tag{4.1.5}$$

This is the first of the two equations which we shall concentrate on below. The quantity

$$R = \gamma s_0 \tag{4.1.6}$$

will below, when we consider the linearized form of (4.1.5) around the uninfected state, be found to equal the mean number of secondary infections produced by one freshly infected individual introduced in a totally susceptible population of size s_0. R is known as the *net reproductive number.* Finally setting

$$x := -\ln(1-p), \qquad p = 1-e^{-x} \tag{4.1.7}$$

gives the *nonlinear renewal equation*

$$x(t) = \gamma s_0 \int_0^t f(x(t-\tau))\psi(\tau)d\tau + h(t) \tag{4.1.8}$$

with

$$f(x) := 1-e^{-x}\,. \tag{4.1.9}$$

With the definitions (4.1.3), (4.1.4), (4.1.7) and (4.1.9) equations (4.1.1), (4.1.2), (4.1.5) and (4.1.8) are all equivalent and we shall freely switch from one to the other depending on which is most convenient for a particular purpose.

EXERCISE 4.1.3: Check the calculations leading from (4.1.1) to (4.1.8). Also follow the route backwards to show the equivalence of all these equations.

As a first step in the analysis of our epidemic model we shall consider its large time behaviour. Since s can only decrease and cannot become negative, $s(t)$ has to go to a limit $s_\infty \geq 0$. It is easier, and more useful, to calculate this limit for the scaled equation (4.1.5). Letting $t \to \infty$ in this equation we obtain

$$-\ln(1-p_\infty) = \gamma s_0 p_\infty + h_\infty\,. \tag{4.1.10}$$

Figure 4.1.1 shows p_∞ for various values of $\pi := f(h_\infty)$.

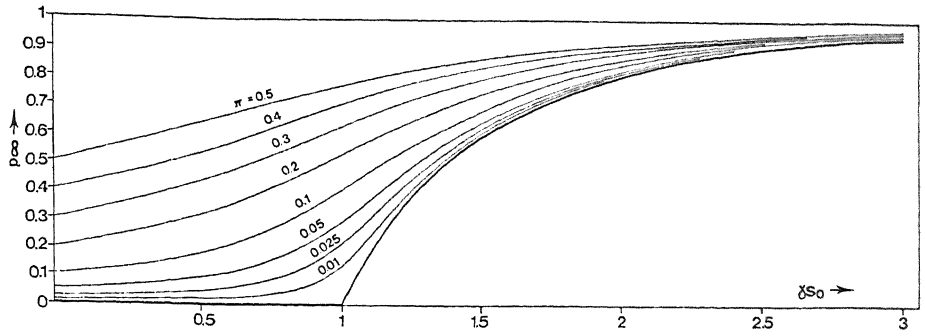

Fig. 4.1.1 The eventual fraction of (new) victims as a function of the number γs_0 of secondary infections produced by one freshly infected individual introduced into a totally uninfected population of size s_0 for various values of $\pi = 1 - e^{-h_\infty}$, where h_∞ is the total starting infectivity.

We see that for small total starting infectivity h_∞ the final epidemic size is negligible for $\gamma s_0 \leqslant 1$ but not for $\gamma s_0 > 1$. This observation is the content of Kermack's and McKendrick's celebrated *threshold theorem:*

$$\lim_{h_\infty \downarrow 0} p_\infty := \hat{p}_\infty \begin{cases} = 0 & \text{for } \gamma s_0 \leqslant 1 \\ > 0 & \text{if } \gamma s_0 > 1 \end{cases}. \tag{4.1.11}$$

The continuity of the solutions of (4.1.1) or its equivalent equations in the forcing functions g or h implies for all fixed $t > 0$ that $p(t) \downarrow 0$ for $h_\infty \downarrow 0$. Hence

$$0 = \lim_{t \uparrow \infty} \lim_{h_\infty \downarrow 0} p(t) \begin{Bmatrix} = \\ < \end{Bmatrix} \lim_{h_\infty \downarrow 0} \lim_{t \uparrow \infty} p(t) = \hat{p}_\infty \begin{cases} \text{for } \gamma s_0 \leqslant 1 \\ \\ \text{for } \gamma s_0 > 1 . \end{cases} \tag{4.1.12}$$

As p necessarily increases monotonically with time, this may be interpreted as a stability of the uninfected state for $\gamma s_0 \leqslant 1$ versus an instability for $\gamma s_0 > 1$.

EXERCISE 4.1.4: Show by graphical means that $\hat{p}_\infty$ is a well-defined nonnegative solution of

$$-\ln(1 - \hat{p}_\infty) = \gamma s_0 \hat{p}_\infty . \tag{4.1.13}$$

Analyse how the solution set of this limit equation depends on γs_0. Which branches are the relevant ones?

EXERCISE 4.1.5: Show that for $\gamma s_0 >> 1$

$$p_\infty = 1 - e^{-(\gamma s_0 + h_\infty)} + o(e^{-(\gamma s_0 + h_\infty)}) \tag{4.1.14}$$

EXERCISE 4.1.6: Write for $\gamma s_0 - 1 << 1$ the relevant solution to (4.1.13) as $\hat{p}_\infty = a_1(\gamma s_0 - 1) + a_2(\gamma s_0 - 1)^2 + \cdots$. Calculate a_1 and a_2. (Hint: For $\gamma s_0 - 1 \gtrless 0$ you have to take different branches of the solution set, and therefore different values of the a_i.)

EXERCISE 4.1.7: Write for fixed $\gamma s_0 < 1$ the solution of (4.1.10) as $p_\infty = a_1 h_\infty + a_2 h_\infty^2 + \cdots$. Calculate a_1 and a_2. Could you extend this procedure to $\gamma s_0 > 1$?

EXERCISE 4.1.8: Analyse the simplest possible of all epidemic models given by

$$\frac{ds}{dt} = -\alpha s y \qquad \frac{dy}{dt} = \alpha s y - \beta y , \tag{4.1.15}$$

y the number of ill individuals. Start with an explicit calculation of the trajectories in the (s,y)-plane followed by an analysis of the equilibria and their stability properties.

EXERCISE 4.1.9: In the model from exercise 4.1.1 let

$$u(t) := \int_0^\infty n(t,\tau)w(\tau)d\tau\,, \qquad w(\tau) := \int_\tau^\infty a(\sigma)\,\Re(\sigma)d\sigma\,/\,\Re(\tau), \qquad \gamma = w(0)\,. \tag{4.1.16}$$

(What is the interpretation of w?) Show that

$$\frac{du}{dt} = \gamma s y - y \tag{4.1.17}$$

with y defined by

$$\frac{ds}{dt} = -ys\,. \tag{4.1.18}$$

Use this result to find a first integral of the p-equations in the form of a relation between s and u. Finally use this first integral to derive Kermack's and McKendrick's threshold theorem for this special case.

EXERCISE 4.1.10: In the model from exercise 4.1.2. let

$$u := WN, \qquad W := -A^T B^{-1}, \qquad \gamma = WC \tag{4.1.19}$$

Same questions as the previous exercise. Hint: The components of $-B^{-1}$ can be interpreted as the mean times that an individual presently in state of illness j will still spend in state of illness i.

So far we concentrated on the fraction of individuals ever to get infected. But we would like to know more. For example, does the maximum number ill at the same time also approach a positive limit for the initial infection going to zero, and if so, how can we characterize this limit? That is, we also want a characterization for small starting infections of the behaviour of the transients from an (almost) fully susceptible p-state to the final p-state after the epidemic has come to its natural end. Such a characterization is provided by what by some authors has been called the "nonlinear renewal theorem" (THIEME, 1985). Here we shall give a heuristic introduction in the context of the epidemic model. Before doing so we shall first have a look at the behaviour of the linearized equation near the noninfected state.

For small p (small x, see (4.1.7)) we have $p \simeq x \simeq f(x)$. Therefore p is approximated by the solution of the linear renewal equation

$$q(t) = \gamma s_0 \int_0^t q(t-\tau)\,\psi(\tau)d\tau + h(t)\,, \tag{4.1.20}$$

as long as both p and q are small. The characteristic equation corresponding to (4.1.20) is

$$1 = \gamma s_0\,\bar{\psi}(r) \tag{4.1.21}$$

Due to our normalization $\bar{\psi}(0)=1$, *i.e.* the net reproductive number R equals γs_0 as anounced previously. So the solution to (4.1.20) will stay bounded when $\gamma s_0<1$ and eventually grow exponentially when $\gamma s_0>1$. This is just the local linearized counterpart of the global stability result derived previously.

EXERCISE 4.1.11: For $\gamma s_0<1$ what tells (4.1.20) you about the behaviour of $q(t)$ for $t\to\infty$? Compare the result with the result from exercise 4.1.7. Hint: remember that $h(t)\to h_\infty$.

Now consider the full equation for $\gamma s_0>1$. If we start with a smaller and smaller starting infection it will take longer and longer before the epidemic reaches an appreciable level. This also means that we stay longer and longer in the region where the linear approximation applies. This in turn means that the detailed form of the initial conditions (the forcing function h) has less and less influence on the behaviour of p when finally the region is reached where the influence of the nonlinearity is felt: by the time p reaches that region the process has already stabilized to exponential growth. But we also have to look further and further into the future to see any interesting behaviour at all. (Fig. 4.1.2 shows an example of this phenomenon.) These observations form the basis for the following:

THEOREM 4.1.12: "Nonlinear renewal theorem, part I". *Assume that* $\gamma s_0>1$. *Choose some* $p_0 \in (0,\hat{p}_\infty)$ *and define* t_0 *by* $p(t_0)=p_0$ *(note that* t_0 *depends on h!). Then for all* t

$$\lim_{h_\infty \downarrow 0} p(t+t_0) = \hat{p}(t) \tag{4.1.22}$$

uniformly in t, *with* $\hat{p}(t)$ *defined as the unique positive solution of*

$$-\ln(1-\hat{p}(t)) = \gamma s_0 \int_0^\infty \hat{p}(t-\tau)\psi(\tau)d\tau \tag{4.1.23}$$

with

$$\hat{p}(0) = p_0 \,. \tag{4.1.24}$$

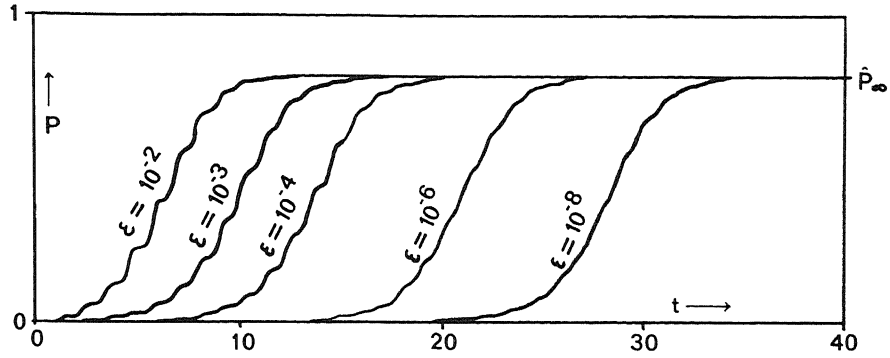

Fig.4.1.2 Fraction of victims as a function of time after an initial infection with a fraction ϵ freshly infected individuals for a scaled infectivity kernel $\psi(\tau)=0$ for $\tau\in[0,1)\cup[1.2,\infty)$ and $\psi(\tau)=5$ for $\tau\in[1,1.2)$, and a threshold parameter $\gamma s_0=2$.

The nonlinear renewal theorem part I is proved in DIEKMANN (1977) for the nonlinear renewal equation (4.1.18) on the general assumption that the function f is sublinear, *i.e.* $f(\alpha x)<\alpha f(x)$ for all $\alpha>0$, but the proof is flawed. DIEKMANN & KAPER (1978) use Tauberian methods to give a rigorous proof of the required estimates in a much more general setting. A completely different approach, using semi-group techniques and the idea of an invariant manifold, is adopted by DIEKMANN & VAN GILS (1984, section 7; for technical simplicity they assume $\psi(\tau)$ to be zero from some τ_0 onwards).

EXERCISE 4.1.13: Verify that f defined by (4.1.9) is sublinear.

* EXERCISE 4.1.14: Find the flaw in DIEKMANN (1977).

The fact that the limit in (4.1.22) is uniform allows us to infer that any quantity like the total number of individuals ill, that can be calculated as a function of time from a convolution of $-\dot{s}=s_0\dot{p}$ with a smooth bounded function h, also converges under the same limiting operation, and that their limits can be calculated from the convolution of h with $s_0\dot{\hat{p}}$. This in turn allows us to infer that also quantities like the maximum number of individuals ever to be ill at the same time converge, thereby answering the question posed earlier.

The "limit epidemics" $\hat{p}$ for different values of p_0 are all translates of each other. They increase monotonically and

$$\lim_{t\uparrow\infty} \hat{p}(t) = \hat{p}_\infty \,. \tag{4.1.25}$$

Moreover

$$\hat{p}(t) = e^{rt}\,(C(p_0) + o(1)) \quad \text{for } t\downarrow-\infty \,, \tag{4.1.26}$$

that is, the behaviour of our limit epidemics for $t\downarrow-\infty$ directly relates to the behaviour of the solution q to our linearized epidemic equation (4.1.20) for $t\uparrow\infty$. This allows us to calculate an estimate for t_0 by an appropriate matching of the two "limits". The result is provided by the following theorem first conjectured by METZ (1978) and subsequently proved by GRIPENBERG (1983c).

THEOREM 4.1.15: "Nonlinear renewal theorem, part II". *Let $\hat{p}_1(t)$ denote the unique solution to (4.1.23) for which*

$$\lim_{t\downarrow-\infty} e^{-rt}\,\hat{p}_1(t) = 1 \,, \tag{4.1.27}$$

then

$$\lim_{h_\infty\downarrow 0} (t_0 + r^{-1}\overline{h}(r)) = \hat{p}_1^{(-1)}(p_0) + r^{-1}\ln\overline{\Phi}(r) \tag{4.1.28a}$$

with

$$\bar{h}(s) = \int_0^\infty e^{-st}\, h(t)dt, \qquad \bar{\Phi}(s) = \int_0^\infty e^{-st}\, t\gamma s_0 \psi(t)dt \tag{4.1.28b}$$

and $\hat{p}_1^{(-1)}:(0,\hat{p}_\infty)\to\mathbb{R}$ *the inverse function of* $\hat{p}_1$.

The two parts of the nonlinear renewal theorem together totally cover the behaviour of the Kermack & McKendrick epidemic for small initial infections (and in practice most initial infections are small!): when h_∞ is small $p(t)$ apparently behaves as some suitable translate of the fixed curve $\hat{p}_1$ independent of the shape of h, where $\hat{p}_1$ is defined by (4.1.23) with (4.1.27) (an effective manner for actually calculating $\hat{p}_1$ is provided by exercise 4.1.17), an estimate of the amount by which $\hat{p}_1$ has to be translated being provided by (4.1.28). As proved, however, the theorem applies to a much wider class of models which can be (re)expressed in the form of the nonlinear renewal equation (4.1.8). See THIEME (1985) for further generalizations.

REMARKS 4.1.16: (i) So far the nonlinear renewal theorem has only been proved under rather restrictive technical assumptions on the function f in (4.1.8). It is to be expected that in the future other versions using other conditions and other convergence concepts will become available *(e.g.* THIEME, 1985).
(ii) METZ (1978) also gives a formulation of the theorem for the full stochastic model with formed the implicit basis of our deterministic equations. This is of some importance as our limiting procedure inevitably brings us into the regime where the number of ill individuals are so low that the law of large numbers on which the deterministic model is essentially based, is bound to fail.

EXERCISE 4.1.17: Calculating the limit epidemic is easiest starting from the scaled limiting form of equation (4.1.1):

$$\dot{\hat{p}}(t) = (1-\hat{p}_1(t))\,\gamma s_0 \int_0^\infty \dot{\hat{p}}_1(t-\tau)\psi(\tau)d\tau\,. \tag{4.1.29}$$

Assume that

$$\hat{p}_1(t) = \sum_{j=1}^{\infty} c_j\, e^{jrt} \tag{4.1.30}$$

By inserting this expression in (4.1.29) derive an equation for r as well as a recurrence relation for the c_j. In METZ (1978) it is shown that for bounded ψ the series expansion (4.1.30) converges on a left half line. Given the value of $\hat{p}_1$ on a left half line we can use any of the equations (4.1.23) or (4.1.29) to extend it numerically to the right.

EXERCISE 4.1.18: Use the result from the previous exercise to find the first few terms in a series expansion of $\ln\hat{p}_1^{(-1)}(p_0)$ for small p_0.

4.2. Population decline in ectotherms

In I.3 a model was presented describing the population growth in ectothermic animals. The main assumptions were that individual growth and reproduction depend on individual size l and food availability x and that the death rate μ was constant except for the possibility of death due to starvation of large animals when their food intake cannot keep pace with their maintenance metabolism. In III.4.3 and III.5.4 this model was extended to include general age dependent mortality as well. The resulting equations were (III.4.3.2):

$$\begin{aligned}
&\frac{\partial l(t,a)}{\partial t} = -\frac{\partial l(t,a)}{\partial a} + g(x,l(t,a)) && 0<a<a_m\leqslant\infty\\
&l(t,0) = l_b\\
&\frac{\partial n(t,a)}{\partial t} = -\frac{\partial n(t,a)}{\partial a} - \mu(a)n(t,a) && \text{for } l(t,a)\leqslant\bar{l}(x)\\
&n(t,a) = 0 && \text{for } l(t,a)>\bar{l}(x)\\
&n(t,0) = b(t)\\
&b(t) = \int_0^{a_m} \lambda(x,l(t,a))n(t,a)da\,.
\end{aligned} \tag{4.2.1a}$$

Equation (4.2.1a) updates the instantaneous age-length relation, (4.2.1b) updates the age distribution and (4.2.1c) is just the birth law. In this section we shall indicate how, given some natural monotonicity assumptions on g and β, the linear theory can give us some easy (but not very spectacular) results about the solutions to (4.2.1) when food availability is dynamically coupled to the population development. (In VI.2 and VI.4 similar results will be presented, but in a different context).

An inspection of the explicit formulae for g, λ and $\bar{l}$ from I.3 tells us that

g is nonnegative, Lipschitz continuous in x and l, increasing in x and decreasing in l, and $g(x,l) = 0$ for $l > \kappa \bar{l}(x)$, $\kappa < 1$,

$\bar{l}$ is nonnegative, and continuous and increasing in x.

λ is bounded, nonnegative, continuous and nondecreasing in x and piecewise continuous in l, and for x sufficiently small $\lambda(x,l) = 0$ for all l.

Moreover we shall assume that

μ is nonnegative, integrable on intervals bounded away from a_m and if $a_m = \infty$ then $\int_0^\infty \mu(a)da = \infty$, *i.e.* individuals cannot live on forever.

First we consider (4.2.1) for x a given, Lipschitz continuous, function of time. In that case (4.2.1a) can be solved simply by integration along the characteristics, giving us $l(t,a)$ for all a and t. Using the known $l(t,a)$ we can then solve (4.2.1b) by integration along the characteristics as

$$n(t,a) = \begin{cases} b(t-a)\exp[-\int_0^a \mu(\alpha)d\alpha]\, \chi(t,0,a) & \text{for } t \geq a \\ n(0,a-t)\exp[-\int_{a-t}^a \mu(\alpha)d\alpha]\, \chi(t,a-t,a) & \text{for } t < a \end{cases} \tag{4.2.2}$$

with

$$\chi(t,a_0,a) := \begin{cases} 1 & \text{if } l(t-a+\alpha,a) \leq \bar{l}(x(t-a+\alpha)) \text{ for all } \alpha \in [a_0,a] \\ 0 & \text{otherwise} \end{cases} \tag{4.2.3}$$

Finally $b(t)$ can be calculated from the usual generation expansion

$$b = \sum_{i=0}^{\infty} b_i \tag{4.2.4a}$$

with

$$b_0(t) = \begin{cases} \int_t^{a_m} \lambda(x(t),l(t,a))n(0,a-t)\exp[-\int_{a-t}^a \mu(\alpha)d\alpha]\, \chi(t,a-t,a)da & \text{for } t \leq a_m \\ 0 & \text{for } t \geq a_m \end{cases} \tag{4.2.4b}$$

$$b_i(t) = \int_0^{\min(t,a_m)} \lambda(x(t),\, l(t,a))b_{i-1}(t-a)\exp[-\int_0^a \mu(\alpha)d\alpha]\, \chi(t,0,a)da\,.$$

In the particular case that x is a fixed constant the death through starvation-mechanism operates at most once, on the initial condition, but after that no individual can grow sufficiently large for the mechanism to come into operation again. Moreover $l(t,a)$ does not depend on t for $a < t$. Therefore b satisfies just the usual linear renewal equation of sections 1 and 2. The various monotonicity properties imply that r is an increasing function of x, and that there exists an x_c such that $r(x_c) = 0$ *(cf.* I.3). Moreover, if $x < x_c$ both $b(t) \to 0$ and $N(t) := \int n(t,a)da \to 0$ for $t \to \infty$, just as in the model with constant death rate discussed in I.3.

Interestingly the last result immediately extends to non constant x: if $x(t) \leq x_c - \epsilon$ for all t then $b(t) \to 0$ and therefore also $N(t) \to 0$ for $t \to \infty$. This can be seen as follows. First we observe that if we shut off the death through starvation mechanism, *i.e.* if we set $\chi = 1$ in (4.2.2) and (4.2.4), this can only make n and b larger. Next we observe that if for the process modified in this manner x' and x'' are two possible courses of the food availability such that $x'(t) \leq x''(t)$ for all t, then also for the corresponding birth rates and age densities $b'(t) \leq b''(t)$ and $n'(a,t) \leq n''(a,t)$

(assuming that the initial condition is the same). Setting $x''(t)=x_c-\epsilon$ then gives the desired result.

Finally we consider the case where the food availability is dynamically coupled to the population:

$$\frac{dx}{dt} = h(x)-u(x)\int_0^{a_m} v(l(t,a))n(t,a)da \tag{4.2.5}$$

with h Lipschitz continuous, positive on $(0,x_e)$ and negative on (x_e,∞), u and v both continuous nonnegative and increasing. The coupled system (4.2.1) and (4.2.5) has at most two equilibra, one in which $N=0$ and $x=x_e$, the other in which $x=x_c$ and N is calculated by setting $dx/dt=0$ in (4.2.5) (compare I.3.5). The latter equilibrium can only exist for $x_c<x_e$. The results derived thus far now allow us to prove that for $x_c>x_e$ the population will go extinct, independent of the initial conditions. To this end we observe that if x satisfies (4.2.5) then $\frac{dx}{dt}<h(x)$. Therefore there exists a t_0 such that $x(t)<x_e=x_c-\epsilon$ for all $t>t_0$. Combining with the result previously obtained for (4.2.1) we arrive at the desired result.

REMARK 4.2.1: It is also possible to extend the argument to cover the case $x_c=x_e$. We shall only give a heuristic sketch. When $x_c=x_e$ there exists a t_0 such that $x(t)<x_c$ for all $t>t_0$. Therefore the mean number of offspring of individuals born after t_0 is bound to be smaller than one. Nonextinction would imply that with the advance of the generations this number converges to one. By (an extension of) the renewal theorem this implies that $b(t)\to\hat{b}\geq 0$. If $\hat{b}$ were positive this would imply that $x(t)<x_c-\epsilon$ from some t onwards for some $\epsilon>0$. But this in turn implies that the mean offspring number does not go to one contradicting the assumption of nonextinction.

5. Models allowing a reduction to a differential equation on $\mathbb{R}^k$.

In this final section we consider two classes of nonlinear age dependent models which allow a reduction of the full p-equation for the age distribution to a differential equation on a finite dimensional space. For the sake of the exposition we shall assume that we are dealing with an autonomous system on the space of age distributions. The extension to systems of p-equations coupled through their inputs and outputs as introduced in III is fairly straightforward. In the first subsection the transformation will be made through a direct application of the linear convergence result from 2.3.1. The second subsection extends the results from 2.3.3 to the nonlinear realm, (but the development is totally independent of that in 2.3.3!) Neither of the two classes of results are specific for age dependent models only. However the restriction to age dependent processes for the second class of problems allows us to give a fairly complete characterization of the nonlinear finite dimensional representability problem.

5.1. Models with a separable death rate

We shall say that a death rate $\mu(a,n)$ depending on age as well as on the p-state n as a whole, is separable when it can be written as

$$\mu(a,n) = \mu_1(a) + \mu_2(n) \tag{5.1.1}$$

If the only nonlinearity in the p-differential generator is through a separable dependence of the death rate on n then it is usually possible to capture the behaviour of the full population model asymptotically for large t by a differential equation in just one variable. This trick seems to be due independently to among others GURTIN, MACCAMY, COLEMAN and SIMMES (see GURTIN, 1980-1982), PRÜSS (1983a), BUSENBERG & IANELLI (1983, 1985) and WEBB (198?a). We shall first illustrate the idea by a simple example, and thereafter put it in a general semigroup setting (thereby effectively lifting the restriction to age dependent models).

EXAMPLE 5.1.1: Consider an age-structured population in which the only density dependence manifests itself through an increase of the age specific death rate $\mu_1(a)$ with an age independent factor $\mu_2(p)$, where p is the weighted average

$$p(t) = \int_0^\infty \gamma(a)n(t,a)da\ . \tag{5.1.2a}$$

The problem

$$\frac{\partial}{\partial t}n(t,a) = -\frac{\partial}{\partial a}n(t,a)-(\mu_1(a)+\mu_2(p))n(t,a) \qquad n(t,0) = \int_0^\infty \beta(a)n(t,a)da \tag{5.1.2b}$$

can be reduced to the linear problem

$$\frac{\partial}{\partial t} m(t,a) = -\frac{\partial}{\partial a} m(t,a) - \mu_1(a)m(t,a) \qquad m(t,0) = \int_0^\infty \beta(a)m(t,a)da \tag{5.1.3}$$

by the (implicit) transformation

$$m(t,a) := e^{\int_0^t \mu_1(\rho(\tau))d\tau} n(t,a) . \tag{5.1.4}$$

The linear theory from 2.3.1 now tells us that

$$m(t,a) = ce^{rt}(\tilde{m}(a) + o(1)) \quad \text{for } t\to\infty , \tag{5.1.5}$$

where r is the dominant real eigenvalue corresponding to (5.1.3) and $\tilde{m}$ is the corresponding stable age distribution. Substitution in (5.1.4) and then into (5.1.2a) yields

$$p(t) = ce^{rt - \int_0^t \mu_1(\rho(\tau))d\tau} [\int_0^\infty \gamma(a)\tilde{m}(a)da + o(1)] . \tag{5.1.6}$$

Omitting the o (1) term we obtain by formal differentiation

$$\frac{dp}{dt} = (r-\mu_2(p))p . \tag{5.1.7}$$

Therefore we conclude that the asymptotic behaviour of n is revealed by (5.1.4), (5.1.5) and a study of (5.1.7).

A rigorous presentation of this useful trick in a more general setting may be found in WEBB (198?a).

A somewhat different approach to the same problem, again in a more general setting, has been taken by BUSENBERG & IANELLI (1983, 1985). They derive a (nonlinear) differential equation for the normalized age distribution

$$w(t,a) := n(t,a) \;/\; \int_0^\infty n(t,\alpha)d\alpha \tag{5.1.8}$$

and observe that one gets one and the same equation for all nonlinearities μ_2, including the special case $\mu_2 = 0$. Hence the normalized age distribution converges to the stable one as $t\to\infty$, leaving an ordinary differential equation problem for the total population size that is asymptotically autonomous.

In conclusion to this subsection we present yet another version of the trick, due to PRÜSS (1983), which closely mimicks the approach taken in II.14. Moreover, we shall, like Webb *(op. cit)*, choose a formulation applying immediately to general structured population problems, not just age dependent ones.

Let A be the generator of a linear semigroup on a Banach space * X, and assume that A has a strictly dominant simple real eigenvalue r with corresponding eigenvector $\tilde{n}$, and that

$$X = \mathcal{N}(A-rI) \oplus \mathcal{R}(A-rI) \tag{5.1.9}$$

where the restriction of the semigroup to the second subspace satisfies an exponential estimate with exponent $r-\epsilon$ for some $\epsilon>0$ Let $\mu_2 : X\to\mathbb{R}$ be smooth. Then the semilinear equation

$$\frac{dn}{dt} = An - \mu_2(n)n \tag{5.1.10}$$

has a solution for each initial condition. We assume that positivity arguments guarantee that the projection onto $\tilde{n}$ according to the decomposition (5.1.9) cannot vanish along a solution if we start with a "positive" initial condition. So for biologically relevant initial data we have the representation

$$n(t,a) = \rho(t)[\tilde{n}(a) + \bar{n}(t,a)] \tag{5.1.11}$$

where $\bar{n}(t,a) \in \mathcal{R}(A-rI)$ and ρ is a real valued function. On substituting (5.1.11) into (5.1.10) and rearranging the terms we find

$$[\frac{d\rho}{dt} - r\rho + \mu_2(n)\rho]\tilde{n} = -\frac{d\rho}{dt}\bar{n} - \rho\frac{d\bar{n}}{dt} + \rho A\bar{n} - \mu_2(n)\rho\bar{n} . \tag{5.1.12}$$

Since $\mathcal{N}(A-rI) \cap \mathcal{R}(A-rI) = \{0\}$ both sides have to be zero. Therefore

---

* See II for the notion of a Banach space and the notation used in formula (5.1.9).

$$\frac{d\rho}{dt} = r\rho - \mu_2(n)\rho \tag{5.1.13}$$

and

$$\frac{d\bar{n}}{dt} = (A - rI)\bar{n} \tag{5.1.14}$$

(Note that (5.1.13) is used to obtain (5.1.14), and that in addition we have divided by ρ.) From (5.1.14) we conclude that

$$\bar{n}(t,\cdot) = 0(e^{-at}) \quad \text{for } t\to\infty\,. \tag{5.1.15}$$

Consequently the asymptotic behaviour of ρ for $t\to\infty$ is determined by the autonomous differential equation in $\mathbb{R}^1$.

$$\frac{d\rho}{dt} = (r - \mu_2(\rho\bar{n}))\rho\,. \tag{5.1.16}$$

* EXERCISE 5.1.2: Formulate hypotheses about μ_2 which allow precise conclusions about the asymptotic behaviour of n. Pay special attention to the case of monotone (with respect to the cone X_+ which is left invariant by the semigroup) μ_2.

5.2. *Linear chain trickery*

In example 1.4.3 and exercise 1.4.4 and in exercise 4.1.2 we have met models which could be phrased as systems of ordinary differential equations, which could be transformed into an age representation. Interestingly enough, under certain conditions we can also effect a reverse transformation thereby providing ourselves with some relatively more easily studied examples of age dependent processes. This idea seems to be due to GURTIN & MACCAMY (1979) and has been used to good avail in diverse practical applications by NISBET & GURNEY and various coworkers (see the contribution by GURNEY, NISBET & BLYTHE in part B for references). Recently it has been discussed in a more abstract setting by HADELER (unpublished manuscript). In this subsection we shall deal with this idea in what we conjecture to be its most general form for age dependent models. Extensions to physiological as opposed to calendar age type models can be found in the contribution by Gurney, Nisbet & Blythe in part B, extensions to size structured models in MURPHY (1983a)).

We start with two examples essentially due to GURTIN & MACCAMY (1979). More examples on different levels of generality may be found in Gurtin's lecture notes (GURTIN, preprint 1982).

EXAMPLE 5.2.1: *A density dependent population model with two individual states.* Consider individuals which can be in two stages, reproductives and seniles, occurring in densities n_1 and n_2 respectively. Furthermore assume that the death rate of both groups equals $\mu(p)$, where p is the total population density, that the birth rates of the reproductive individuals equals $\beta(p)$ and that reproductive individuals become senile at a constant rate α, then

$$\frac{dn_1}{dt} = \beta(p)n_1 - \alpha n_1 - \mu(p)n_1\,, \tag{5.2.1a}$$

$$\frac{dn_2}{dt} = \alpha n_1 - \mu(p)n_2\,, \tag{5.2.1b}$$

$$p = n_1 + n_2\,. \tag{5.2.1c}$$

Alternatively this equation may be rewritten in terms of the population density and the density of reproductives

$$\frac{dn_1}{dt} = \beta(p)n_1 - \alpha n_1 - \mu(p)n_1\,, \tag{5.2.2a}$$

$$\frac{dp}{dt} = \beta(p)n_1 - \mu(p)p\,, \tag{5.2.2b}$$

where (5.2.2b) is obtained by adding (5.2.1a) and (5.2.1b).
The model embodied in equation (5.2.1) or (5.2.2) was specifically designed to allow an age representation: all individuals are born equal, the progression through the stages is independent of outside influences and the death mechanism

---

* We shall in this subsection revert to the notational convention of III in that we shall drop without warning arguments to which no direct reference is made.

is nonselective. To derive an equation for the age distribution we first observe that the death rate is independent of the stage an individual is in and therefore independent of age, so *

$$\frac{\partial n}{\partial t} = -\frac{\partial n}{\partial a} - \mu(p)n \,, \qquad n(t,0) = b(t) \,, \tag{5.2.3a}$$

where p is just the total density of individuals

$$p(t) = \int_0^\infty n(t,a)da \,. \tag{5.2.3b}$$

To calculate the birth rate b we note that the probability that an individual aged a is still in the reproductive state equals $e^{-\alpha a}$, so

$$b(t) = \beta(p) \int_0^\infty e^{-\alpha a}\, n(t,a)da \,. \tag{5.2.3c}$$

Only calculating *the* initial age distribution presents some difficulty if we have only $p(0)$ and $n_1(0)$ to go by. However, for practical purposes we may take just any age distribution compatible with there being $n_1(0)$ reproductives and $n_2(0)$ seniles at $t=0$ *i.e.* any $n(0,\cdot)$ satisfying

$$\int_0^\infty n(0,a)da = p(0) \,, \qquad \int_0^\infty e^{-\alpha a}\, n(0,a)da = n_1(0) \,. \tag{5.2.4}$$

Finally we shall effect the backward transformation from equation (5.2.3) into equation (5.2.2) again. The clue is provided by equation (5.2.4): what will do for the initial population state should do just as well for the population state at any later time. Therefore, for any age dependent process satisfying (5.2.3) we just define

$$n_1(t) := \int_0^\infty e^{-\alpha a}\, n(t,a)da \,, \qquad p(t) := \int_0^\infty n(t,a)da \,. \tag{5.2.5}$$

Then

$$\frac{dn_1}{dt} = \frac{d}{dt}\int_0^\infty n(t,a)e^{-\alpha a}\,da = \int_0^\infty \frac{\partial n(t,a)}{\partial t}\, e^{-\alpha a}\,da = -\int_0^\infty [\frac{\partial n(t,a)}{\partial a} + \mu(p)n(t,a)]e^{-\alpha a}\,da =$$

$$-\, n(t,a)e^{-\alpha a}\Big|_0^\infty - \alpha \int_0^\infty n(t,a)e^{-\alpha a}\,da - \mu(p)n_1 = \beta(p)n_1 - \alpha n_1 - \mu(p)n_1 \,, \tag{5.2.6a}$$

where the last step is effected by using the side condition (5.2.3c) and analogously we find

$$\frac{dp}{dt} = \beta(p)n_1 - \mu(p)\,p \,. \tag{5.2.6b}$$

EXAMPLE 5.2.2: *A density dependent population model with three individual states.* This example is exactly equal to the previous one, except that we assume three stages: juvenile, reproductive adult, and senile, with equal transition rates between stages 1 and 2 and stages 2 and 3. The stage based equations are

$$\begin{aligned} \frac{dn_1}{dt} &= \beta(p)n_2 - \alpha n_1 - \mu(p)n_1 \,, \\ \frac{dn_2}{dt} &= \alpha n_1 - \alpha n_2 - \mu(p)n_2 \,, \\ \frac{dp}{dt} &= \beta(p)n_2 - \mu(p)p \,, \end{aligned} \tag{5.2.7}$$

with corresponding age based equations

$$\begin{aligned} &\frac{\partial n}{\partial t} = -\frac{\partial n}{\partial a} - \mu(p)n \,, \quad n(t,0) = b(t) \,, \\ &p(t) = \int_0^\infty n(t,a)da \,, \qquad b(t) = \beta(p)\int_0^\infty \alpha a e^{-\alpha a} n(t,a)da \,, \end{aligned} \tag{5.2.8}$$

and connecting rules

$$n_1(t) = \int_0^\infty e^{-\alpha a} n(t,a)da, \qquad n_2(t) = \int_0^\infty \alpha a e^{-\alpha a} n(t,a)da, \qquad p(t) = \int_0^\infty n(t,a)da \,. \tag{5.2.9}$$

EXERCISE 5.2.3: Check the transformation from (5.2.7) to (5.2.8) vice versa.

The stage models introduced in the examples are of no great interest in themselves, but the backward transformation rules are: suppose we were given some model which eventually led to equation (5.2.3) or equation (5.2.8), then it would always be possible to effect the transformation to equations (5.2.2) or (5.2.7) irrespective of the specific origin of the model. If we consider a particular age dependent mechanism like *e.g.* cannibalism, but we have some freedom of choice still in how we let the various quantities involved depend on the age distribution, then we could at least try whether a model of a form like (5.2.3) or (5.2.8) would be in our class, in order to find a relatively easily studied example. This then leaves us with two questions: (1) what does the ordinary differential equation counterpart tell us about the full partial differential equation, and (2) how far can the trick be extended?

Suppose we are given a partial differential equation of the form

$$\frac{\partial n(t,a)}{\partial t} = -\frac{\partial n(t,a)}{\partial a} - \mu(a,n)n(t,a), \quad n(t,a) = b(t) = \beta(n) \tag{5.2.10}$$

with $\mu(a,\cdot)$ and β functionals on the space of age distributions X. Equation (5.2.10) is supposed to define a dynamical system on X. Abstractly it can be rewritten as

$$\frac{dn}{dt} = F(n)\,. \tag{5.2.11}$$

The procedure in the examples then was (compare (5.2.5) and (5.2.6)) to seek a continuous linear map $P:X\to\mathbb{R}^k$ sending n into N such that

$$\frac{dN}{dt} = \frac{d}{dt}Pn = P\frac{dn}{dt} = PF(n) = G(Pn) = G(N) \tag{5.2.12}$$

for some function $G:\mathbb{R}^k\to\mathbb{R}^k$. If this procedure is really good it should also allow us to calculate $n(t,a)$ again at least for sufficiently large t. Integrating (5.2.10) along the characteristics shows that $n(t,a)$ for $a<t$ is fully determined by the solution of (5.2.12) if and only if both μ and β are so determined, *i.e.* it should be possible to write $\mu(a,n)=\mu_0(a,N)$ and $\beta(n)=\beta_0(N)$. This answers our first question, and also part of our second one.

The remaining part of the answer to the second question should be provided by analyzing what restrictions on μ_0, β_0 and P can be deduced from (5.2.12). As a first step we write the linear map P as

$$Pn = \int_0^\infty \Phi(a)n(a)da\,. \tag{5.2.13}$$

(Mathematically this implies that we have to make some assumptions about the space X.) Using (5.2.13) we find

$$PF(n) = \int_0^\infty \Phi(a)\left(-\frac{\partial n}{\partial a}(a) - \mu_0(a,Pn)n(a)\right)da = \Phi(0)\beta_o(Pn) + \int_0^\infty \left(\frac{d\Phi}{da}(a) - \Phi(a)\mu_0(a,Pn)\right)n(a)da\,, \tag{5.2.14}$$

(using that $\Phi(a)n(a)$ should go to zero for $a\to\infty$ from (5.2.13)). Inserting (5.2.14) in (5.2.12) gives

$$\int_0^\infty \left(\frac{d\Phi}{da}(a) - \mu_0(a,Pn)\Phi(a)\right)n(a)da = H(Pn) := G(Pn) - \Phi(0)\beta_0(Pn)\,. \tag{5.2.15}$$

Taking (Fréchet) derivatives for n at both sides and setting

$$\mu_1(a) := \mu_0(0,a), \qquad B := DH(0) \tag{5.2.16}$$

gives

$$\int_0^\infty \left(\frac{d\Phi}{da}(a) - \mu_1(a)\Phi(a)\right)n(a)da = B\int_0^\infty \Phi(a)n(a)da\,. \tag{5.2.17}$$

(5.2.17) can only be true for all n when

$$\frac{d\Phi}{da} = (B+\mu_1(a)I)\Phi(a)\,, \tag{5.2.18}$$

from which we conclude that

$$\Phi(a) = (\mathfrak{F}(a))^{-1}e^{Ba}C \tag{5.2.19a}$$

with

$$C := \Phi(0), \qquad \mathfrak{F}(a) := e^{-\int_0^a \mu_1(\alpha)d\alpha}\,. \tag{5.2.19b}$$

Finally inserting (5.2.18) in (5.2.15) gives.

$$\int_0^\infty (\mu_1(a)-\mu_0(a,Pn))\Phi(a)n(a)da = H(Pn)-BPn . \tag{5.2.20}$$

The right hand side of (5.2.20) is a function of Pn. Therefore the left hand side should also depend on n through Pn only. A sufficient (and we expect also necessary) condition for this is that

$$\mu_0(a,Pn) = \mu_1(a) + \mu_2(Pn) , \tag{5.2.21}$$

which is nothing but our old friend from 5.1, the separability condition.

Summarizing, we have found that if in (5.2.10)

$$\beta(n) = \beta_0(Pn), \qquad \mu(a,n) = \mu_1(a) + \mu_2(Pn) \tag{5.2.22}$$

where the linear map P is defined by

$$Pn = \int_0^\infty (\mathfrak{F}(a))^{-1} e^{Ba} Cn(a)da \tag{5.2.23}$$

with $\mathfrak{F}$ defined by (5.2.19b), then $N:=Pn$ satisfies

$$\frac{dN}{dt} = BN + C\beta_0(N)-\mu_2(N)N . \tag{5.2.24}$$

We also conjecture that this statement can be reversed: if there exists a continuous linear map $P:X\to\mathbb{R}^k$ such that (5.2.10) holds with $\beta(n)=\beta_0(Pn)$ and $\mu(a,n)=\mu_0(a,Pn)$ and $N:=Pn$ satisfies a differential equation $dN/dt=G(N)$, then P is necessarily of the form (5.2.23), μ of the form (5.2.22) and G of the form (5.2.24).

EXERCISE 5.2.4: Let B be the differential generator of a Markov chain with killing, *i.e.* $b_{ij}\geqslant 0$ for $i\neq j$, $b_{ii}\leqslant 0$ and $B^TE + U=0$ with all $u_i\geqslant 0$ and $E=(1,..,1)^T$, and let C be a probability vector, *i.e.* all $c_i\geqslant 0$ and $C^TE=1$. Now consider the population model

$$\frac{dN}{dt} = BN + C\beta_0(N)-\mu_2(N)N . \tag{5.2.25}$$

Derive the corresponding age representation and also the back transformation to (5.2.25) again.

EXERCISE 5.2.4: Let in (5.2.10)

$$\beta(n) = \frac{\theta\int_0^\infty \alpha_1 a e^{-\alpha_1 a} n(a)da}{1+\int_0^\infty n(a)da}$$

and

$$\mu(n,a) = \gamma_1+\gamma_2\int_0^\infty e^{\alpha_2 a} n(a)da .$$

Find a k and a linear map $P:X\to\mathbb{R}^k$ such that $\beta(n)=\beta_0(Pn)$, $\mu(n,a) = \mu_2(Pn)$ and $N=Pn$ satisfies $dN/dt=G(N)$ for some function G.

We shall end this section with a few words about the practical implications of the result embodied in formula (5.2.22) to (5.2.24). Generally a model specification will begin with a specification of the functionals β and μ. If our reduction procedure is to apply, then it should be possible to write $\mu(a,n)$ as $\mu_1(a)+\mu_3(n)$. From μ_1 we can calculate $\mathfrak{F}$ using its definition (5.2.19b). Next we observe that the components of $e^{Ba}C$ are necessarily all mixtures of polynomials times exponentials times sines and cosines. Conversely the realization theorems from systems theory (see e.g. KALMAN, FALB & ARBIB, 1969; BROCKETT, 1970; SILVERMAN, 1971; PADULO & ARBIB, 1974 or Wsillems, 1975) imply that any set of functions $\psi_i(a)$, $i=1,..h$ of that form can be expressed as $A_ie^{Ba}C$ for some $k\times k-$ matrix B and some vectors A_i and C. So we can find a representation of the form (5.2.24) if and only if $\beta(n)$ and $\mu_3(n)$ can both be calculated from some finite number of integrals $\int_0^\infty \psi_i(a)(\mathfrak{F}(a))^{-1}n(a)da$ with ψ_i of the indicated form. For practical purposes it is moreover preferrable to choose k as small as possible. In most cases we dream up as examples, finding such a minimal representation of the $\psi_i(a)$ will not present a great problem. A guideline of how to proceed in general

can be found in the references on systems realization mentioned earlier.

* EXERCISE 5.2.5: Discuss how the results from this subsection relate to those presented in 2.3.3.

** EXERCISE 5.2.6: Either prove that (5.2.21) is also a necessary condition for (5.2.20) to hold true for all n, or give a counter example.

A. The Laplace transformation

This appendix surveys the bare essentials about the Laplace transformation for non-cognoscenti. A detailed account may be found *e.g.* in WIDDER (1946), DOETSCH (1950, 1955, 1956) and VAN DER POL & BREMMER (1955). An extensive set of tables is Erdélyi *et al.* (1954).

Suppose that $f:\mathbb{R}^+ \to \mathbb{R}$ is some sufficiently well-behaved function, then its Laplace transform $\bar{f}$ is defined as

$$\bar{f}(s) := \int_0^\infty e^{-st} f(t)dt \tag{A.1}$$

for those complex s for which the integral converges. It can be shown that convergence of the integral for some $s_0=\sigma_0+i\tau_0$ entails the convergence for all $s=\sigma+i\tau$ with $\sigma>\sigma_0$, and moreover $\bar{f}$ is analytic in the half plane $\mathrm{Re}(s)>\sigma_0$. Generally we shall call by the same name the Laplace transform as defined by (A.1) and the analytic extension of $\bar{f}$ as a function of s. Finally, if f satisfies some exponential bound, $|f(t)| < ce^{\lambda t}$, then the integral certainly converges for $\mathrm{Re}(s)>\lambda$. These considerations are enough to justify many of our formal manipulations. The main usefulness of the Laplace transform, however, derives from the facts that 1) generally speaking the relation between f and $\bar{f}$ is one to one, and 2) we can deduce properties of f from $\bar{f}$ and vice versa.

Some examples of Laplace transform pairs are*

exponential density

$$f(t) = \alpha e^{-\alpha t} \quad \Leftrightarrow \bar{f}(s) = \frac{\alpha}{\alpha+s} \tag{A.2}$$

gamma density

$$f(t) = \frac{\alpha(\alpha t)^{k-1}e^{-\alpha t}}{\Gamma(k)} \quad \Leftrightarrow \bar{f}(s) = \left[\frac{\alpha}{\alpha+s}\right]^k \tag{A.3}$$

homogeneous density

$$f(t) = \begin{cases} (\beta-\alpha)^{-1} & \text{for } \alpha<t<\beta \\ 0 & \text{elsewhere} \end{cases} \Leftrightarrow \bar{f}(s) = \frac{e^{-\alpha s}-e^{-\beta s}}{(\beta-\alpha)s} \tag{A.4}$$

and in the limit, the delta function

$$f(t) = \delta(t-\alpha) \quad \Leftrightarrow \quad \bar{f}(s) = e^{-\alpha s} \tag{A.5}$$

Here we have chosen our f such that they integrate to one, *i.e.* $\bar{f}(0) = 1$. This is no restriction due to the linearity of the Laplace transform

$$f = \alpha f_1 + \beta f_2 \quad \Leftrightarrow \quad \bar{f} = \alpha\bar{f}_1 + \beta\bar{f}_2 \tag{A.6}$$

A class of functions which are not integrable but which allow a Laplace transform for $\mathrm{Re}(s)>0$ are

$$f(t) = t^\alpha,\ \alpha>-1 \quad \Leftrightarrow \quad \bar{f}(s) = \frac{\Gamma(\alpha+1)}{s^{\alpha+1}} \tag{A.7}$$

Finally two immensely useful properties of the Laplace transform, one in the context of integral equations, the other in the context of differential equations, are

$$f = f_1 * f_2 \quad \Leftrightarrow \quad \bar{f} = \bar{f}_1\bar{f}_2 \tag{A.8}$$

$$f_2 = \frac{df_1}{dt} \quad \Leftrightarrow \bar{f}_2 = s\bar{f}_1 - \bar{f}_1(0)\,. \tag{A.9}$$

---

* The gamma function, Γ, is defined as $\Gamma(x) := \int_0^\infty t^{x-1}e^{-t}dt$. For integer x this reduces to (as can be shown by partial integration) $\Gamma(n) = (n-1)!$

EXERCISE A.1: Deduce the explicit solution (2.1.10) by Laplace transforming both sides of the renewal equation (2.1.7).

EXERCISE A.2: Prove (A.2) to (A.9). Hint: For (A.9) use partial integration.

Using (A.2) to (A.9) we can derive all other sorts of useful relations like

$$f_0(t) = \int_0^t f_1(\tau)d\tau \quad \Leftrightarrow \quad \bar{f}_0(s) = \bar{f}_1(s)/s \tag{A.10}$$

from (A.7) and (A.8) with $f_2(t) = t^0$,

$$f(t) = \int_t^\infty f_1(\tau)d\tau \quad \Leftrightarrow \quad \bar{f}(s) = (\int_0^\infty f_1(t)dt - \bar{f}_1(s))/s \tag{A.11}$$

from (A.10) and (A.6), and

$$f(t) = \begin{cases} f_1(t-\alpha) & \text{for } t>\alpha \\ 0 & \text{for } t<\alpha \end{cases} \Leftrightarrow \bar{f}(s) = e^{-\alpha s}\bar{f}_1(s) \tag{A.12}$$

from (A.5) and (A.8) with $f_2(t)=\delta(t-\alpha)$.

The problem of finding f from $\bar{f}$ is called the inversion problem. The crucial result is that for most classes of functions, in particular the continuous ones, the function f is uniquely determined by $\bar{f}$, with the understanding that two functions that differ only on a set of measure zero are counted as equivalent. Therefore it is possible to "calculate" some unknown f from its known Laplace transform $\bar{f}$ by looking $\bar{f}$ up, possibly after some rearrangement, in one of the extensive published tables of Laplace transform pairs. However, more often than not it is not possible for a given $\bar{f}$ to arrive at a pleasing formula for f. The main use of Laplace transform methods then is to deduce from $\bar{f}$ results about the behaviour of f for large t. (For example, the so-called Tauberian theorems relate the behaviour of f for large t to that of $\bar{f}$ near $s=0$.) Therefore Laplace transform methods are in a sense complementary to numerical methods which mainly concentrate on short time behaviour.

Various explicit formulae for f in terms of $\bar{f}$ are known, the most useful one being

$$f(t) = \frac{1}{2\pi i}\int_{\alpha-i\infty}^{\alpha+i\infty} e^{st}\bar{f}(s)ds \tag{A.13}$$

α being chosen such that all singularities of $\bar{f}$ lie to the left of the vertical line of integration $\mathrm{Re}(s)=\alpha$. The main utility of (A.13) comes from the fact that it allows us to use so-called contour integration to extract information about f from $\bar{f}$. We shall conclude this appendix with two examples in which we use contour integration to find (more or less) explicit expressions for f.

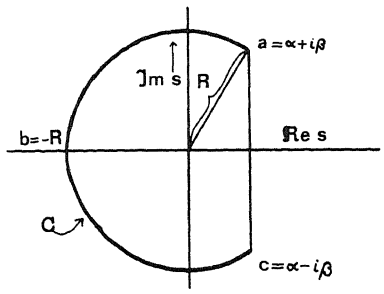

Fig. A.1. Integration contour used to evaluate (A.13).

EXAMPLE A.3: If $\bar{f}$ is meromorphic, *i.e.* its only singularities are poles, and if for the contour C depicted in fig. A.1

$$|\bar{f}(s)| < MR^{-\epsilon},\ M,\epsilon>0 \text{ for } s \text{ on the arc abc of } C \tag{A.14}$$

then for $R\to\infty$ the integral of $e^{st}\bar{f}(s)$ on the arc *abc* goes to zero, *i.e.* the integral around the contour C becomes equal to (A.13). But by the so-called residue theorem

$$\frac{1}{2\pi i}\oint_C e^{st}\bar{f}(s)da = \sum \text{ residues of } e^{st}\bar{f}(s) \text{ at poles inside } C \tag{A.15a}$$

where the residue of an n'th order pole at s_j equals

$$\lim_{s\to s_j} \frac{1}{(n-1)!} \frac{d^{n-1}}{ds^{n-1}} (s-s_j)^n e^{st}\bar{f}(s) = p(t)e^{s_j t} \tag{A.15b}$$

with p a polynomial of degree $n-1$.

The principal class of Laplace transforms satisfying (A.15) are the rational functions $\bar{f}(s)=p_1(s)/p_2(s)$ with the degree of the polynomial p_2 larger than that of p_1. In renewal theory the Laplace transform of the birth rate function

$$\bar{b} = \frac{\bar{g}}{1-\bar{\phi}} \tag{A.16}$$

is of this form if both ϕ and g are mixtures of polynomials times (possibly complex) exponentials like ϕ and g from example 1.4.3.

EXERCISE A.4: Calculate b for $\phi(a)=R\alpha^2 a e^{-\alpha a}$ and $g(t) = \gamma e^{\beta t}$

EXAMPLE A.4: When $\bar{f}$ is meromorphic but has infinetely many poles then (A.14) can never hold true. However, if we can find a sequence $R_i\to\infty$ such that (A.14) holds for those R_i, then we can still calculate f by taking the limit in (A.15). As an example of this procedure consider the renewal equation with $\phi(a)=R\beta^{-1}$ for $a\leqslant\beta$, $\phi(a)=0$ for $a>\beta$. Then

$$\bar{b}(s) = \frac{\beta s\bar{g}(s)}{\beta s - R + Re^{-\beta s}} . \tag{A.17}$$

If $g(t)=0$ for $t>\beta$ then $\bar{g}$ has no poles. Therefore the poles of $\bar{b}$ coincide with the solutions of

$$\frac{\beta s - R + \mathrm{Re}^{-\beta s}}{\beta s} = 0 \tag{A.18}$$

This equation has but one real solution r and all other solutions s_j have smaller real part. Moreover all the s_j are simple zeros of $(\beta s)^{-1}(\beta s - R + \mathrm{Re}^{-\beta s})$, so that all poles of $\bar{b}$ are of order 1. If we number the roots in order of decreasing real part, roots with positive imaginary part taking precedence over roots with negative imaginary part, then for $R\neq 1$ the results in BELLMAN & COOKE (1963) imply that for $t>\beta$

$$b(t) = \sum_{j=0}^{\infty} \frac{s_j\bar{g}(s_j)}{1-Re^{-\beta s_j}} e^{s_j t} \tag{A.19}$$