# Life insurance mathematics

**Abstract:** The subject matter and methodology of modern life insurance mathematics are surveyed. Standard insurance products with payments depending only on life history events are described and analyzed in the commonly used Markov chain model under the assumption of deterministic interest rates. The actuarial equivalence principle for determining premiums and reserves is motivated by risk diversification in large portfolios. Differential equations for reserves and other conditional expected values are obtained and proclaimed basic constructive and computational tools. The traditional method for managing non-diversifiable risk arising from changes in e.g. mortality rates and interest rates is outlined; premiums are calculated under prudent assumptions, and systematic surpluses thus created are repaid as bonus. An alternative way of managing non-diversifiable risk is to link the contractual payments to the development of mortality, interest, and possibly other demographic or economic indexes. Examples of so-called index-linked products are given, and premiums and reserves are determined by combined use of classical actuarial principles and principles of pricing and hedging from financial mathematics. Transfer of non-diversifiable risk to the financial markets through creation of tradeable insurance derivatives is outlined as an idea. Methods for estimation of mortality and other vital rates, formerly a major issue in life insurance mathematics, are fetched from modern statistical life history analysis and therefore only briefly described.

*Key-words:* Life annuity, life assurance, life endowment, life history analysis, time-continuous Markov chain, semi-Markov chain, compound interest, principle of equivalence, arbitrage pricing theory, with profit contract, surplus and bonus, defined benefits, defined contributions, index linked insurance, securitization.

**A. The scope of this survey.** Fetching its tools from a number of disciplines, actuarial science is defined by the direction of its applications rather than by the models and methods it makes use of. However, there exists at least one stream of actuarial science that presents problems and methods sufficiently distinct to merit status as an independent area of scientific inquiry – Life insurance mathematics. This is the area of applied mathematics that studies risk associated with life and pension insurance and methods for management of such risk. Under this strict definition, its models and methods are described as they present themselves today, sustained by contemporary mathematics, statistics, and computation technology. No attempt is made to survey its two millenniums of past history, its vast collection of techniques that were needed before the advent of scientific computation, or the countless details that arise from its practical implementation to a wide range of products in a strictly regulated industry.

**B. Basic notions of payments and interest.** In the general context of financial services, consider a financial contract between a company and a customer. The terms and conditions of the contract generate a stream of payments, benefits from the company to the customer less contributions from the customer to the company. They are represented by a payment function

$$B(t) = \text{total amount paid in the time interval [0,t]}\,,$$

$t \geq 0$, time being reckoned from the inception of the policy. This function is taken to be right-continuous, $B(t) = \lim_{\tau \searrow t} B(\tau)$, with existing left-limits, $B(t-) = \lim_{\tau \nearrow t} B(\tau)$. When different from 0, the jump $\Delta B(t) = B(t) - B(t-)$ represents a lump sum payment at time $t$. There may also be payments due continuously at rate $b(t)$ per time unit at any time $t$. Thus, the payments made in any small time interval $[t, t + dt)$ total

$$dB(t) = b(t)\,dt + \Delta B(t)\,.$$

Payments are currently invested (contributions deposited and benefits withdrawn) into an account that bears interest at rate $r(t)$ at time $t$. The balance of the company's account at time $t$, called the *retrospective reserve*, is the sum of all past and present payments accumulated with interest,

$$\mathcal{U}(t) = \int_{0-}^{t} e^{\int_\tau^t r(s)\,ds} d(-B)(\tau)\,,$$

where $\int_{0-}^{t} = \int_{[0,\,t]}$. Upon differentiating this relationship, one obtains the dynamics

$$d\mathcal{U}(t) = \mathcal{U}(t)\,r(t)\,dt - dB(t)\,, \tag{1}$$

showing how the balance increases with interest earned on the current reserve and net payments into the account.

The company's discounted (strictly) future net liability at time $t$ is

$$\mathcal{V}(t) = \int_{t}^{n} e^{-\int_t^\tau r(s)\,ds} dB(\tau)\,. \tag{2}$$

Its dynamics,

$$d\mathcal{V}(t) = \mathcal{V}(t)\, r(t)\, dt - dB(t)\,,$$

shows how the debt increases with interest and decreases with net redemption.

**C. Valuation of financial contracts.** At any time $t$ the company has to provide a *prospective reserve* $V(t)$ that adequately meets its future obligations under the contract. If the future payments and interest rates would be known at time $t$ (e.g. a fixed plan savings account in an economy with fixed interest), then so would the present value in (2), and an adequate reserve would be $V(t) = \mathcal{V}(t)$. For the contract to be economically feasible, no party should profit at the expense of the other, so the value of the contributions must be equal to the value of the benefits at the outset:

$$\Delta B(0) + V(0) = 0\,. \tag{3}$$

If future payments and interest rates are uncertain so that $\mathcal{V}(t)$ is unknown at time $t$, then reserving must involve some principles beyond those of mere accounting. A clear-cut case is when the payments are *derivatives* (functions) of certain assets (securities), one of which is a money account with interest rate $r$. Principles for valuation of such contracts are delivered by modern financial mathematics, see e.g. [3]. We will describe them briefly in lay terms. Suppose there is a finite number of basic assets that can be traded freely, in unlimited positive or negative amounts (taking long or short positions) and without transaction costs. An investment *portfolio* is given by the number of shares held in each asset at any time. The size and the composition of the portfolio may be dynamically changed through sales and purchases of shares. The portfolio is said to be *self-financing* if, after its initiation, every purchase of shares in some assets is fully financed by selling shares in some other assets (the portfolio is currently rebalanced, with no further infusion or withdrawal of capital). A self-financing portfolio is an *arbitrage* if the initial investment is negative (the investor borrows the money) and the value of the portfolio at some later time is non-negative with probability one. A fundamental requirement for a market to be well-functioning is that it does not admit arbitrage opportunities. A financial claim that depends entirely on the prices of the basic securities is called a financial *derivative*. Such a claim is said to be *attainable* if it can be perfectly reproduced by a self-financing portfolio. The no arbitrage regime dictates that the price of a financial derivative must at any time be the current value of the self-financing portfolio that reproduces it. It turns out that the price at time $t$ of a stream of attainable derivative payments $B$ is of the form

$$V(t) = \tilde{\mathbb{E}}\left[\mathcal{V}(t) \middle| \mathcal{G}_t\right] , \tag{4}$$

where $\tilde{\mathbb{E}}$ is the expected value under a so-called *equivalent martingale measure* derived from the price processes of the basic assets, and $\mathcal{G}_t$ represent all market events observed up to and including time $t$. Again, the contract must satisfy (3), which determines the initial investment $-\Delta B(0)$ needed to buy the self-financing portfolio. The financial market is *complete* if every financial derivative is attainable. If the market is not complete, there is no unique price for every derivative, but any pricing principle obeying the no arbitrage requirement must be of the form (4).

**D. Valuation of life insurance contracts by the principle of equivalence.** Characteristic features of life insurance contracts are, firstly, that the payments are contingent on uncertain individual life history events (largely unrelated to market events) and, secondly, that the contracts are long term and binding to the insurer. Therefore, there exists no liquid market for such contracts and they can not be valued by the mere principles of accounting and finance described above. Pricing of life insurance contracts and management of the risk associated with them are the paramount issues of life insurance mathematics.

The paradigm of traditional life insurance is the so-called *principle of equivalence*, which is based on the notion of risk diversification in a large portfolio under the assumption that the future development of interest rates and other relevant economic and demographic conditions are known. Consider a portfolio of $m$ contracts currently in force. Switching to calendar time, denote the payment function of contract No $i$ by $B^i$ and denote its discounted future net liabilities at time $t$ by $\mathcal{V}^i(t) = \int_t^n e^{-\int_t^\tau r(s)\,ds} dB^i(\tau)$. Let $\mathcal{H}_t$ represent the life history by time $t$ (all individual life history events observed up to and including time $t$) and introduce $\xi_i = \mathbb{E}[\mathcal{V}^i(t)|\mathcal{H}_t]$, $\sigma_i^2 = \mathbb{V}\mathrm{ar}[\mathcal{V}^i(t)|\mathcal{H}_t]$, and $s_m^2 = \sum_{i=1}^m \sigma_i^2$. Assume that the contracts are independent and that the portfolio grows in a balanced manner such that $s_m^2$ goes to infinity without being dominated by any single term $\sigma_i^2$. Then, by the central limit theorem, the standardized sum of total discounted liabilities converges in law to the standard normal distribution as the size of the portfolio increases:

$$\frac{\sum_{i=1}^m (\mathcal{V}^i(t) - \xi_i)}{s_m} \xrightarrow{\mathcal{L}} \mathrm{N}(0,1)\,. \tag{5}$$

Suppose the company provides individual reserves given by

$$V^i(t) = \xi_i + \epsilon\, \sigma_i^2 \tag{6}$$

for some $\epsilon > 0$. Then

$$\mathbb{P}\left[ \sum_{i=1}^{m} V^i(t) - \sum_{i=1}^{m} \mathcal{V}^i(t) > 0 \,\middle|\, \mathcal{H}_t \right] = \mathbb{P}\left[ \frac{\sum_{i=1}^{m}(\mathcal{V}^i(t) - \xi_i)}{s_m} < \epsilon\, s_m \,\middle|\, \mathcal{H}_t \right],$$

and it follows from (5) that the total reserve covers the discounted liabilities with a (conditional) probability that tends to 1. Similarly, taking $\epsilon < 0$ in (6), the total reserve covers discounted liabilities with a probability that tends to 0. The benchmark value $\epsilon = 0$ defines the actuarial *principle of equivalence*, which for an individual contract reads (drop topscript $i$):

$$V(t) = \mathbb{E}\left[ \int_t^n e^{-\int_t^\tau r(s)\, ds} dB(\tau) \,\middle|\, \mathcal{H}_t \right]. \tag{7}$$

In particular, for given benefits the premiums should be designed so as to satisfy (3).

**E. Life and pension insurance products.** Consider a life insurance policy issued at time 0 for a finite term of $n$ years. There is a finite set of possible states of the policy, $\mathcal{Z} = \{0, 1, \ldots, J\}$, 0 being the initial state. Denote the state of the policy at time $t$ by $Z(t)$. The uncertain course of policy is modeled by taking $Z$ to be a stochastic process. Regarded as a function from $[0, n]$ to $\mathcal{Z}$, $Z$ is assumed to be right-continuous, with a finite number of jumps, and commencing from $Z(0) = 0$. We associate with the process $Z$ the *indicator processes* $I_g$ and *counting processes* $N_{gh}$ defined, respectively, by $I_g(t) = 1[Z(t) = g]$ (1 or 0 according as the policy is in the state $g$ or not at time $t$) and $N_{gh}(t) = \sharp\{\tau;\, Z(\tau-) = g,\, Z(\tau) = h,\ \tau \in (0, t]\}$ (the number of transitions from state $g$ to state $h$ $(h \neq g)$ during the time interval $(0, t]$).

The payments $B$ generated by an insurance policy are typically of the form

$$dB(t) = \sum_g I_g(t)\, dB_g(t) + \sum_{g \neq h} b_{gh}(t)\, dN_{gh}(t)\,, \tag{8}$$

where each $B_g$ is a payment function specifying payments due during sojourns in state $g$ (a general *life annuity*), and each $b_{gh}$ specifying lump sum payments due upon transitions from state $g$ to state $h$ (a general *life assurance*). When different from 0, $\Delta B_g(t)$ represents a lump sum (general *life endowment*) payable in state $g$ at time $t$. Positive amounts represent benefits and negative amounts represent premiums. In practice premiums are only of annuity type.

Figure 1 shows a flow-chart for a policy on a single life with payments dependent only on survival and death. We list the basic forms of benefits: An $n$-year *term insurance* with sum insured 1 payable immediately upon death, $b_{01}(t) = 1[t \in (0,n)]$; An $n$-year life endowment with sum 1, $\Delta B_0(n) = 1$; An $n$-year life annuity payable annually in arrears, $\Delta B_0(t) = 1$, $t = 1,\ldots,n$; An $n$-year life annuity payable continuously at rate 1 per year, $b_0(t) = 1[t \in (0,n)]$; An $(n-m)$-year annuity deferred in $m$ years payable continuously at rate 1 per year, $b_0(t) = 1[t \in (m,n)]$. Thus, an $n$-year term insurance with sum insured $b$ against premium payable continuously at rate $c$ per year is given by $dB(t) = b\,dN_{01}(t) - c\,I_0(t)\,dt$ for $0 \le t < n$ and $dB(t) = 0$ for $t \ge n$.

The flow-chart in Figure 2 is apt to describe a single-life policy with payments that may depend on the state of health of the insured. For instance, an $n$-year *endowment insurance* (a combined term insurance and and life endowment) with sum insured $b$, against premium payable continuously at rate $c$ while active (waiver of premium during disability), is given by $dB(t) = b\,(dN_{02}(t) + dN_{12}(t)) - c\,I_0(t)\,dt$, $0 \le t < n$, $dB(n) = b\,(I_0(n) + I_1(n))$, $dB(t) = 0$ for $t > n$.

The flow-chart in Figure 3 is apt to describe a multi-life policy involving three lives called $x$, $y$, and $z$. For instance, an $n$-year insurance with sum $b$ payable upon the death of the last survivor against premium payable as long as all three are alive is given by $dB(t) = b\,(dN_{47}(t) + dN_{57}(t) + dN_{67}(t)) - c\,I_0(t)\,dt$, $0 \le t < n$, $dB(t) = 0$ for $t \ge n$.

**F. The Markov chain model for the policy history.** The breakthrough of stochastic processes in life insurance mathematics was marked by Hoem's 1969 paper [10], where the process $Z$ was modeled as a time-continuous Markov chain. The Markov property means that the future course of the process is independent of its past if the present state is known: For $0 < t_1 < \cdots < t_q$ and $j_1,\ldots,j_q$ in $\mathcal{Z}$,

$$\mathbb{P}[Z(t_q) = j_q \,|\, Z(t_p) = j_p\,, p = 1,\ldots,q-1] = \mathbb{P}[Z(t_q) = j_q \,|\, Z(t_{q-1}) = j_{q-1}]\,.$$

It follows that the *simple transition probabilities*,

$$p_{gh}(t,u) = \mathbb{P}[Z(u) = h \,|\, Z(t) = g]\,,$$

determine the finite-dimensional marginal distributions through

$$\mathbb{P}[Z(t_p) = j_p\,, p = 1, \ldots, q] = p_{0j_1}(0, t_1)\, p_{j_1j_2}(t_1, t_2) \cdots p_{j_{q-1}j_q}(t_{q-1}, t_q)\,,$$

hence they also determine the entire probability law of the process $Z$. It is moreover assumed that, for each pair of states $g \neq h$ and each time $t$, the limit

$$\mu_{gh}(t) = \lim_{u \searrow t} \frac{p_{gh}(t,u)}{u-t}\,,$$

exists. It is called the *intensity of transition* from state $g$ to state $h$ at time $t$. In other words,

$$p_{gh}(t,u) \;=\; \mu_{gh}(t)\, dt \,+\, o(dt)\,,$$

where $o(dt)$ denotes a term such that $o(dt)/dt \to 0$ as $dt \to 0$. The intensities, being one-dimensional and easy to interpret as "instantaneous conditional probabilities of transition per time unit", are the basic entities in the probability model. They determine the simple transition probabilities uniquely as solutions to sets of differential equations. The *Kolmogorov backward differential equations* for the $p_{jg}(t,u)$, seen as functions of $t \in [0,u]$ for fixed $g$ and $u$, are

$$\frac{\partial}{\partial t} p_{jg}(t,u) = -\sum_{k;k\neq j} \mu_{jk}(t)\,(p_{kg}(t,u) - p_{jg}(t,u))\,, \tag{9}$$

with side conditions $p_{gg}(u,u) = 1$ and $p_{jg}(u,u) = 0$ for $j \neq g$. The *Kolmogorov forward equations* for the $p_{gj}(s,t)$, seen as functions of $t \in [s,n]$ for fixed $g$ and $s$, are

$$\frac{\partial}{\partial t} p_{gj}(s,t) = \sum_{i;i\neq j} p_{gi}(s,t)\mu_{ij}(t) \;- p_{gj}(s,t)\mu_{j\cdot}(t)\,, \tag{10}$$

with obvious side conditions at $t = s$. The forward equations are sometimes the more convenient because, for any fixed $t$, the functions $p_{gj}(s,t)$, $j = 0, \ldots, J$, are probabilities of disjoint events and therefore sum to 1. A technique for obtaining such differential equations is sketched in Paragraph G.

A differential equation for the *sojourn probability*,

$$p_{\overline{gg}}(t,u) = \mathbb{P}[Z(\tau) = g\,, \tau \in (t,u] \,|\, Z(t) = g]\,,$$

is easily put up and solved to give

$$p_{\overline{gg}}(t,u) = e^{-\int_t^u \mu_{g\,\cdot}(s)\, ds}\,,$$

where $\mu_{g\,\cdot}(t) = \sum_{h,\,h\neq g} \mu_{gh}(t)$ is the *total intensity of transition* out of state $g$ at time $t$.

To see that the intensities govern the probability law of the process $Z$, consider a fully specified path of $Z$, starting from the initial state $g_0 = 0$ at time $t_0 = 0$, sojourning there until time $t_1$, making a transition from $g_0$ to $g_1$ in $[t_1, t_1 + dt_1)$, sojourning there until time $t_2$, making a transition from $g_1$ to $g_2$ in $[t_2, t_2 + dt_2)$, and so on until making its final transition from $g_{q-2}$ to $g_{q-1}$ in $[t_{q-1}, t_{q-1} + dt_{q-1})$, and sojourning there until time $t_q = n$. The probability of this elementary event is a product of sojourn probabilities and infinitesimal transition probabilities, hence a function only of the intensities:

$$e^{-\int_{t_0}^{t_1} \mu_{g_0\cdot}(s)\,ds}\,\mu_{g_0 g_1}(t_1)\,dt_1\,e^{-\int_{t_1}^{t_2} \mu_{g_1\cdot}(s)\,ds}\,\mu_{g_1 g_2}(t_2)\,dt_2 \cdots e^{-\int_{t_{q-1}}^{t_q} \mu_{g_{q-1}\cdot}(s)\,ds}$$

$$= e^{\sum_{p=1}^{q-1} \ln \mu_{g_{p-1}g_p}(t_p) - \sum_{p=1}^{q} \sum_{h;h\neq g_{p-1}} \int_{t_{p-1}}^{t_p} \mu_{g_{p-1}h}(s)\,ds}\,dt_1 \ldots dt_{q-1}\,. \quad (11)$$

**G. Actuarial analysis of standard insurance products.** The bulk of existing life insurance mathematics deals with the situation where the functions $B_g$ and $b_{gh}$ depend only on the life history of the individual(s) covered under the policy. We will be referring to such products as *standard*. Moreover, interest rates and intensities of transition are assumed to be deterministic (known at time 0).

We consider first simple products with payments $dB_g(t)$ and $b_{gh}(t)$ depending only on the policy duration $t$ (as the notation indicates). Then, with “memoryless” payments and policy process (the Markov assumption), the reserve in (7) is a function of the time $t$ and the current policy state $Z(t)$ only. Therefore, we need only determine the state-wise reserves

$$V_j(t) = \mathbb{E}\left[\int_t^n e^{-\int_t^\tau r(s)\,ds} dB(\tau)\,\middle|\, Z(t) = j\right]. \quad (12)$$

Inserting (8) into (12) and using the obvious relationships

$$\mathbb{E}\,[I_g(\tau)\,|\,Z(t) = j] = p_{jg}(t,\tau)\,,$$

$$\mathbb{E}\,[dN_{gh}(\tau)\,|\,Z(t) = j] = p_{jg}(t,\tau)\,\mu_{gh}(\tau)\,d\tau\,,$$

we obtain

$$V_j(t) = \int_t^n e^{-\int_t^\tau r(s)\,ds} \sum_g p_{jg}(t,\tau) \left( dB_g(\tau) + \sum_{h;h\neq g} \mu_{gh}(\tau)\,b_{gh}(\tau)\,d\tau \right). \quad (13)$$

It is an almost universal principle in continuous time stochastic processes theory that conditional expected values of functions of the future, given the past, are solutions to certain differential equations. More often than not these are needed to construct the solution. Therefore, the theory of differential equations and numerical methods for solving them are part and parcel of stochastic processes and their applications.

The state-wise reserves $V_j$ satisfy the first order ordinary differential equations (ODE)

$$\frac{d}{dt}V_j(t) = r(t)\,V_j(t) - b_j(t) - \sum_{k;k\neq j} \mu_{jk}(t)\,(b_{jk}(t) + V_k(t) - V_j(t))\,, \qquad (14)$$

valid at all times $t$ where the coefficients $r$, $\mu_{jk}$, $b_j$, and $b_{jk}$ are continuous and there are no lump sum annuity payments. The ultimo conditions

$$V_j(n-) = \Delta B_j(n)\,, \qquad (15)$$

$j = 1, \ldots, J$, follow from the very definition of the reserve. Likewise, at times $t$ where annuity lump sums are due,

$$V_j(t-) = \Delta B_j(t) + V_j(t)\,. \qquad (16)$$

The equations (14) are so-called *backward* differential equations since the solution is to be computed backwards starting from (15).

The differential equations can be derived in various ways. We will sketch a simple heuristic method called *direct backward construction*, which works due to the piece-wise deterministic behaviour of the Markov chain. Split the expression on the right of (2) into

$$\mathcal{V}(t) = dB(t) + e^{-r(t)\,dt}\,\mathcal{V}(t + dt)$$

(suppressing a negligible term $o(dt)$) and condition on what happens in the time interval $(t, t + dt]$. With probability $1 - \mu_{j\,\cdot}(t)\,dt$ the policy stays in state $j$ and, conditional on this, $dB(t) = b_j(t)\,dt$ and the expected value of $\mathcal{V}(t + dt)$ is $V_j(t + dt)$. With probability $\mu_{jk}(t)\,dt$ the policy moves to state $j$ and, conditional on this, $dB(t) = b_{jk}(t)$ and the expected value of value of $\mathcal{V}(t + dt)$ is $V_k(t + dt)$. One gathers

$$\begin{aligned} V_j(t) \;&=\; (1 - \mu_{j\,\cdot}(t)\,dt)\left(b_j(t)\,dt + e^{-r(t)\,dt}\,V_j(t + dt)\right) \\ &\quad + \sum_{k;\,k\neq j} \mu_{jk}(t)\,dt\left(b_{jk}(t) + e^{-r(t)\,dt}\,V_k(t + dt)\right) + o(dt)\,. \qquad (17) \end{aligned}$$

Rearranging, dividing by $dt$, and letting $dt \to 0$, one arrives at (14).

In the single life model sketched in Figure 1, consider an endowment insurance with sum $b$ against premium at level rate $c$ under constant interest rate $r$. The differential equation for $V_0$ is

$$\frac{d}{dt} V_0(t) \quad = \quad r\, V_0(t) + c - \mu(t)\,(b - V_0(t))\,, \tag{18}$$

subject to $V_0(n-) = b$. This is *Thiele's differential equation* discovered in 1875.

The expression on the right of (14) shows how the reserve, seen as a debt, increases with interest (first term) and decreases with redemption of annuity type in the current state (second term) and of lump sum type upon transition to other states (third term). The quantity

$$R_{jk} = b_{jk}(t) + V_k(t) - V_j(t) \tag{19}$$

appearing in the third term is called the *sum at risk* in respect of transition from state $j$ to state $k$ at time $t$ since it is the amount credited to the insured's account upon such a transition: the lump sum payable immediately plus the adjustment of the reserve. This sum multiplied with the rate $\mu_{jk}(t)$ is a rate of expected payments.

Solving (14) with respect to $-b_j(t)$, which can be seen as a premium (rate), shows that the premium consists of a *savings premium* $\frac{d}{dt}V_j(t) - r(t)\,V_j(t)$ needed to maintain the reserve (the increase of the reserve less the interest it earns) and a *risk premium* $\sum_{k;k\neq j} \mu_{jk}(t)\,(b_{jk}(t) + V_k(t) - V_j(t))$ needed to cover risk due to transitions.

The differential equations (14) are as transparent as the defining integral expressions (13) themselves, but there are other and more important reasons why they are useful.

Firstly, the easiest (and sometimes the only) way of computing the values of the reserves is by solving the differential equations numerically (e.g by some finite difference method). The coefficients in the equations are precisely the elemental functions that are specified in the model and in the contract. Thus all values of the state-wise reserves are obtained in one run. The integrals (13) might be computed numerically, but that would require separate computation of the transition probabilities as functions of $\tau$ for each given $t$. In general the transition probabilities are themselves compound quantities that can only be obtained as solutions to differential equations.

Secondly, the differential equations are indispensable constructive tools when more complicated products are considered. For instance, if the life endowment contract behind (18) is modified such that 50% of the reserve is paid out upon death in addition to the sum insured, then its differential equation becomes

$$\frac{d}{dt}V_0(t) \quad = \quad r\,V_0(t) + c - \mu(t)\,(b - 0.5V_0(t))\,, \tag{20}$$

which is just as easy as (18).

Another point in case are administration expenses, which are treated as benefits and covered by charging the policyholder an extra premium in accordance with the equivalence principle. Such expenses may incur upon the inception of the policy (included in $\Delta B_0(0)$), as annuity type payments (included in the $b_g(t)$), and in connection with payments of death and endowment benefits (included in the $b_{gh}(t)$ and the $\Delta B_g(t)$). In particular, expenses related to the company's investment operations are typically allocated to the individual policies on a pro rata basis, in proportion to their individual reserves. Thus, for our generic policy, there is a cost element running at rate $\gamma(t)\,V_j(t)$ at time $t$ in state $j$. Subtracting this term on the right hand side of (14) creates no difficulty and, virtually, is just to the effect of reducing the interest rate $r(t)$.

The non-central conditional moments

$$V_j^{(q)}(t) = \mathbb{E}[\mathcal{V}(t)^q | Z(t) = j]\,,$$

$q = 1, 2, \ldots$, do not in general possess explicit integral expressions. They are, however, solutions to the backward differential equations

$$\begin{aligned}\frac{d}{dt}V_j^{(q)}(t) \quad &= \quad (qr(t) + \mu_{j\cdot}(t))V_j^{(q)}(t) \; - \; q\,b_j(t)V_j^{(q-1)}(t) \\ &\quad - \sum_{k;\,k\neq j} \mu_{jk}(t) \sum_{p=0}^{q} \binom{q}{p} (b_{jk}(t))^p V_k^{(q-p)}(t)\,,\end{aligned}$$

subject to the conditions $V_j^{(q)}(n-) = \Delta B_j(n)^q$ (plus joining conditions at times with annuity lump sums). The backward argument goes as for the reserves, only with a few more details to attend to, starting from

$$\begin{aligned}\mathcal{V}(t)^q \quad &= \quad \Big(dB(t) + e^{-r(t)\,dt}\mathcal{V}(t+dt)\Big)^{q-p} \\ &= \quad \sum_{p=0}^{q} \binom{q}{p} dB(t)^q\, e^{-r(t)\,dt\,(q-p)}\, \mathcal{V}(t+dt)^{q-p}\,.\end{aligned}$$

Higher order moments shed light on the risk associated with the portfolio. Recalling (5) and using the notation in Paragraph D, a solvency margin approximately equal to the upper $\varepsilon$-fractile in the distribution of the discounted outstanding net liability is given by $\sum_{i=1}^{m} \xi_i + c_{1-\varepsilon} s_m$, where $c_{1-\varepsilon}$ is the upper $\varepsilon$-fractile of the standard normal distribution. More refined estimates of the fractiles of the total liability can be obtained by involving three or more moments.

**H. Path-dependent payments and semi-Markov models.** The technical matters in Paragraphs F-G become more involved if the contractual payments or the transition intensities depend on the past life history. We will consider examples where they may depend on the sojourn time $S(t)$ that has elapsed since entry into the current state, henceforth called the *state duration.* Thus, if $Z(t) = j$ and $S(t-) = s$ at policy duration $t$, the transition intensities and payments are of the form $\mu_{jk}(s,t)$, $b_j(s,t)$, and $b_{jk}(s,t)$. To ease exposition, we disregard intermediate lump sum annuities, but allow of terminal endowments $\Delta B_j(s,n)$ at time $n$. The state-wise reserve will now be a function of the form $V_j(s,t)$.

In simple situations (e.g. no possibility of return to previously visited states) one may still work out integral expressions for probabilities and reserves (they will be multiple integrals). The differential equation approach always works, however. The relationship (17) modifies to

$$\begin{aligned} V_j(s,t) &= (1 - \mu_{j\,\cdot}(s,t)\,dt)\left(b_j(s,t)\,dt + e^{-r(t)\,dt} V_j(s+dt, t+dt)\right) \\ &\quad + \sum_{k;\,k\neq j} \mu_{jk}(s,t)\,dt\ \left(b_{jk}(s,t) + V_k(0,t)\right) + o(dt)\,, \end{aligned}$$

from which one obtains the first order partial differential equations

$$\begin{aligned} \frac{\partial}{\partial t} V_j(s,t) &= r(t) V_j(s,t) - \frac{\partial}{\partial s} V_j(s,t) - b_j(s,t) \\ &\quad - \sum_{k;k\neq j} \mu_{jk}(s,t)\left(b_{jk}(s,t) + V_k(0,t) - V_j(s,t)\right)\,, \end{aligned} \tag{21}$$

subject to the conditions

$$V_j(s, n-) = \Delta B_j(s,n)\,. \tag{22}$$

We give two examples of payments dependent on state duration. In the framework of the disability model in Figure 2, an $n$-year disability annuity payable at rate 1 only after a qualifying period of $q$, is given by $b_1(s,t) =$

$1[q < s < t < n]$. In the framework of the three-lives model sketched in Figure 3, an $n$-year term insurance of 1 payable upon the death of $y$ if $z$ is still alive and $x$ is dead and has been so for at least $q$ years, is given by $b_{14}(s,t) = 1[q < s < t < n]$.

Probability models with intensities dependent on state duration are known as a semi-Markov models. A case of support to their relevance is the disability model in Figure 2. If there are various forms of disability, then the state duration may carry information about the severity of the disability and hence about prospects of longevity and recovery.

**I. Managing non-diversifiable risk for standard insurance products.** Life insurance policies are typically long term contracts, with time horizons wide enough to see substantial variations in interest, mortality, and other economic and demographic conditions affecting the economic result of the portfolio. The rigid conditions of the standard contract leave no room for the insurer to meet adverse developments of such conditions; he can not cancel contracts that are in force and also not reduce their benefits or raise their premiums. Therefore, with standard insurance products there is associated a risk that cannot be diversified by increasing the size of the portfolio as in Paragraph D. The limit operation leading to (7) was made under the assumption of fixed interest. In an extended set-up, with random economic and demographic factors, this amounts to conditioning on $\mathcal{G}_n$, the economic and demographic development over the term of the contract. Instead of (7), one gets

$$V(t) = \mathbb{E}\left[ \int_t^n e^{-\int_t^\tau r(s)\,ds} dB(\tau) \middle| \mathcal{H}_t,\, \mathcal{G}_n \right] . \tag{23}$$

At time $t$ only $\mathcal{G}_t$ is known, so (23) is not a feasible reserve. In particular, the equivalence principle (3), recast as

$$\Delta B_0(0) + \mathbb{E}\left[ \int_0^n e^{-\int_0^\tau r(s)\,ds} dB(\tau) \middle| \mathcal{G}_n \right] = 0\,, \tag{24}$$

is also infeasible since benefits and premiums are fixed at time 0 when $\mathcal{G}_n$ cannot be anticipated.

The traditional way of managing the non-diversifiable risk is to charge premiums sufficiently high to cover, on the average in the portfolio, the contractual benefits under all likely economic-demographic scenarios. The systematic surpluses that (most likely) will be generated by such prudently calculated premiums belong to the insured and are paid back in arrears as the history $\mathcal{G}_t$ unfolds. Such contracts are called *participating policies* or *with-profit contracts*.

The repayments, called *dividends* or *bonus*, are represented by a payment function $D$. They should be controlled in such a manner as to ultimately restore equivalence when the full history is known at time $n$:

$$\Delta B_0(0) + \mathbb{E}\left[\int_0^n e^{-\int_0^\tau r(s)\,ds}(dB(\tau) + dD(\tau))\middle|\,\mathcal{G}_n\right] = 0\,, \tag{25}$$

A common way of designing the prudent premium plan is, in the framework of Paragraphs D-H, to calculate premiums and reserves on a so-called *technical basis* with interest rate $r^*$ and transition intensities $\mu^*_{jk}$ that represent a worst-case-scenario. Equipping all technical quantities with an asterisk, we denote the corresponding reserves by $V^*_j$, the sums at risk by $R^*_{jk}$ etc. The surplus generated by time $t$ is, quite naturally, defined as the excess of the factual retrospective reserve over the contractual prospective reserve,

$$\mathcal{S}(t) = \mathcal{U}(t) - \sum_{j=0}^{J} I_j\, V^*_j(t)\,. \tag{26}$$

Upon differentiating this expression, using (1), (8), (14), and the obvious relationship $dI_j(t) = \sum_{k;\,k\neq j}(dN_{kj}(t) - dN_{jk}(t))$, one obtains after some rearrangement that

$$d\mathcal{S}(t) = \mathcal{S}(t)\, r(t)\, dt + dC(t) + dM(t)\,, \tag{27}$$

where

$$dC(t) = \sum_{j=0}^{J} I_j(t)\, c_j(t)\, dt\,, \tag{28}$$

$$c_j(t) = (r(t) - r^*)\, V^*_j(t) + \sum_{k;\,k\neq j} R^*_{jk}(t)(\mu^*_{jk}(t) - \mu_{jk}(t))\,, \tag{29}$$

$$dM(t) = -\sum_{j\neq k} R^*_{jk}(dN_{jk}(t) - I_j(t)\,\mu_{jk}(t)\,dt)\,. \tag{30}$$

The right hand side of (27) displays the dynamics of the surplus. The first term is the interest earned on the current surplus. The last term, given by (30), is purely erratic and represents the policy's instantaneous random deviation from the expected development. The second term, given by (28), is the systematic contribution to surplus, and (29) shows how it decomposes into gains due to prudent assumptions about interest and transition intensities in each state.

One may show that (25) is equivalent to

$$\mathbb{E}\left[\int_0^n e^{-\int_0^\tau r(s)\,ds}(dC(\tau) - dD(\tau))\,\middle|\, \mathcal{G}_n\right] = 0\,, \tag{31}$$

which says that, on the average over the portfolio, all surpluses are to be repaid as dividends.

The dividend payments $D$ are controlled by the insurer. Since negative dividends are not allowed, it is possible that (31) cannot be ultimately attained if the contributions to surplus turn negative (the technical basis was not sufficiently prudent) and/or dividends were paid out prematurely. Designing the dividend plan $D$ is therefore a major issue, and it can be seen as a problem in optimal stochastic control theory. For a general account of the point process version of this theory, see [5]. Various schemes used in practice are described in [15]. The simplest (and least cautious one) is the so-called *contribution plan*, whereby surpluses are repaid currently as they arise: $D = C$.

The much discussed issue of *guaranteed interest* takes a clear form in the framework of the present theory. Focusing on interest, suppose the technical intensities $\mu^*_{jk}$ are the same as the factual $\mu_{jk}$ so that the surplus emerges only from the excess of the factual interest rate $r(t)$ over the technical rate $r^*$:

$$dC(t) = (r(t) - r^*)\,V^*_{Z(t)}(t)\,dt\,.$$

Under the contribution plan $dD(t)$ must be set to 0 if $dC(t) < 0$, and the insurer will therefore have to cover the negative contributions $dC^-(t) = (r^* - r(t))_+\,V^*_{Z(t)}(t)\,dt$. Averaging out the life history randomness, the discounted value of these claims is

$$\int_0^n e^{-\int_0^\tau r(s)\,ds}(r^* - r(\tau))_+ \sum_{j=0}^J p_{0j}(0,\tau)\,V^*_j(\tau)\,d\tau\,. \tag{32}$$

Mobilizing the principles of arbitrage pricing theory set out in Paragraph C, we conclude that the interest guarantee inherent in the present scheme has a market price which is the expected value of (32) under the equivalent martingale measure. Charging a down premium equal to this price at time 0 would eliminate the downside risk of the contribution plan without violating the format of the with-profit scheme.

**J. Unit-linked insurance.** The dividends $D$ redistributed to holders of standard with-profit contracts can be seen as a way of adapting the benefits payments to the development of the non-diversifiable risk factors of $\mathcal{G}_t$, $0 < t \leq n$. Alternatively one could specify in the very terms of the contract that the payments will depend, not only on life history events, but also on the development of interest, mortality, and other economic-demographic conditions. One such approach is the unit-linked contract, which relates the benefits to the performance of the insurer's investment portfolio. To keep things simple, let the interest rate $r(t)$ be the only uncertain non-diversifiable factor. A straightforward way of eliminating the interest rate risk is to let the payment function under the contract be of the form

$$dB(t) = e^{\int_0^t r(s)\, ds}\, dB^0(t)\,, \tag{33}$$

where $B^0$ is a baseline payment function dependent only on the life history. This means that all payments, premiums and benefits, are index-regulated with the value of a unit of the investment portfolio. Inserting this into (24), assuming that life history events and market events are independent, the equivalence requirement becomes

$$\Delta B^0_{\scriptscriptstyle 0}(0) + \mathbb{E}\left[\int_0^n dB^0(\tau)\right] = 0\,.$$

This requirement does not involve the future interest rates and can be met by setting an equivalence baseline premium level at time 0.

Perfect unit-linked products of the form (33) are not offered in practice. Typically, only the sum insured (of e.g. a term insurance or a life endowment) is index-regulated while the premiums are not. Moreover, the contract usually comes with a guarantee that the sum insured will not be less than a certain nominal amount. Averaging out over the life histories, the payments become purely financial derivatives, and pricing goes by the principles described in Paragraph C. If random life history events are kept part of the model, one faces a pricing problem in an incomplete market. This problem was formulated and solved in [13] in the framework of the theory of risk minimization [7].

**K. Defined benefits and defined contributions.** With-profit and unit-linked are two just two ways of adapting benefits to the long-term development of non-diversifiable risk factors. The former does not include the adaptation rule in the terms of the contract, whereas the latter does. We mention two other arch-type insurance schemes that are widely used in practice. *Defined benefits* means literally that only the benefits are specified in the contract, either in nominal figures as in the with-profit contract or in units of some index. A commonly used index is the salary (final or average) of the insured. In that case also the contributions (premiums) are usually linked to the salary (typically a certain percentage of the annual income). Risk management of such a scheme is a matter of designing the rule for collection of contributions. Unless the future benefits can be precisely predicted or reproduced by dynamic investment portfolios, defined benefits leave the insurer with a major non-diversifiable risk. Defined benefits are gradually being replaced with their opposite, *defined contributions*, with only premiums specified in the contract. This scheme has much in common with the traditional with-profit scheme, but leaves more flexibility to the insurer as benefits do not come with a minimum guarantee.

**L. Securitization.** Generally speaking, and recalling Paragraph C, any introduction of new securities in a market helps to complete it. *Securitization* means creating tradeable securities that may serve to make non-traded claims attainable. This device, well known and widely used in the commodities markets, was introduced in non-life insurance in the 1990-es when exchanges and insurance corporations launched various forms of *insurance derivatives* aimed to transfer catastrophe risk to the financial markets. Securitization of non-diversifiable risk in life insurance, e.g. through bonds with coupons related to mortality experience, is conceivable. If successful, it would open new opportunities of financial management of life insurance risk by the principles described in Paragraph C. A work in this spirit is [16], where market attitudes are modeled for all forms of risk associated with a life insurance portfolio, leading to market values for reserves.

**M. Statistical inference.** The theory of inference in point process models is a well developed area of statistical science, see e.g. [1]. We will just indicate how it applies to the Markov chain model and only consider the parametric case where the intensities are of the form $\mu_{gh}(t;\theta)$ with $\theta$ some finite-dimensional parameter.

The likelihood function for an individual policy is obtained upon inserting the observed policy history (the processes $I_g$ and $N_{gh}$) in (11) and dropping the $dt_i$:

$$\Lambda = \exp\left(\sum_{g\neq h}\int(\ln\mu_{gh}(\tau)dN_{gh}(\tau)-\mu_{gh}(\tau)I_g(\tau)d\tau)\right). \tag{34}$$

The integral ranges over the period of observation. In this context the time parameter $t$ will typically be the age of the insured.

The total likelihood for the observations from a portfolio of $m$ independent risk is the product of the individual likelihoods and, therefore, of the same form as (34), with $I_g$ and $N_{gh}$ replaced by $\sum_{i=1}^m I_g^i$ and $\sum_{i=1}^m N_{gh}^i$. The maximum likelihood estimator $\hat{\theta}$ of the parameter vector $\theta$ is obtained as the solution to the likelihood equations

$$\frac{\partial}{\partial\theta}\ln\Lambda\bigg|_{\theta=\hat{\theta}} = 0\,.$$

Under regularity conditions, $\hat{\theta}$ is asymptotically normally distributed with mean $\theta$ and a variance matrix that is the inverse of the information matrix

$$\mathbb{E}_\theta\left(-\frac{\partial^2}{\partial\theta\partial\theta'}\ln\Lambda\right).$$

This result forms the basis for construction of tests and confidence intervals.

A technique that is specifically actuarial, starts from the provisional assumption that the intensities are piece-wise constant, e.g. $\mu_{gh}(t)=\mu_{gh;j}$ for $t\in[j-1,j)$, and that the $\mu_{gh;j}$ are functionally unrelated and thus constitute the entries in $\theta$. The maximum likelihoood estimators are then the so-called *occurrence-exposure rates*

$$\hat{\mu}_{gh;j} = \frac{\int_{j-1}^{j}\sum_{i=1}^m dN_{gh}^i(\tau)}{\int_{j-1}^{j}\sum_{i=1}^m I_g^i d\tau}\,,$$

which are empirical counterparts to the intensities. A second stage in the procedure consists in fitting parametric functions to the occurrence-exposure rates by some technique of (usually non-linear) regression. In actuarial terminology this is called *analytic graduation.*

**N. A remark on notions of reserves.** Retrospective and prospective reserves were defined in [14] as conditional expected values of $\mathcal{U}(t)$ and $\mathcal{V}(t)$, respectively, given some information $\mathcal{H}_t'$ available at time $t$. The notions of reserves used here conform with that definition, taking $\mathcal{H}'_t$ to be the full information $\mathcal{H}_t$. While the definition of the prospective reserve never was a matter of dispute in life insurance mathematics, there exists an alternative notion of retrospective reserve which we shall describe. Under the hypothesis of deterministic interest, the principle of equivalence (3) can be recast as

$$\mathbb{E}[\mathcal{U}(t)] = \mathbb{E}[\mathcal{V}(t)]\,. \tag{35}$$

In the single life model this reduces to $\mathbb{E}[\mathcal{U}(t)] = p_{00}(0,t)V_0(t)$, hence

$$V_0(t) = \frac{\mathbb{E}[\mathcal{U}(t)]}{p_{00}(0,t)}\,.$$

This expression, expounded as "the fund per survivor", was traditionally called the *retrospective reserve.* A more descriptive name would be *the retrospective formula for the prospective reserve (under the principle of equivalence).* For a multi-state policy (35) assumes the form

$$\mathbb{E}[\mathcal{U}(t)] = \sum_g p_{0g}(0,t)\, V_g(t)\,.$$

As this is only one constraint on $J$ functions, (35) alone does not provide a non-ambiguous notion of state-wise retrospective reserves.

**O. A view to the literature.** In this brief survey of contemporary life insurance mathematics no space has been left to the wealth of techniques that now have mainly historical interest, and no attempt has been made to trace the origins of modern ideas and results. References to the literature are selected accordingly, their purpose being to add details to the picture drawn here with broad strokes of the brush. A key reference on the early history of life insurance mathematics is [9]. Textbooks covering classical insurance mathematics are [2], [4], [6], [8], and [12]. An account of counting processes and martingale techniques in life insurance mathematics can be compiled from [11] and [15].

# References


[1] Andersen, P.K., Borgan, Ø., Gill, R.D., Keiding, N. (1993). *Statistical Models Based on Counting Processes.* Springer-Verlag.

[2] Berger, A. (1939): *Mathematik der Lebensversicherung.* Verlag von Julius Springer, Vienna.

[3] Björk, T. (1998): *Arbitrage Theory in Continuous Time.* Oxford University Press.

[4] Bowers, N.L. Jr., Gerber, H.U., Hickman, J.C., and Nesbitt, C.J. (1986). *Actuarial Mathematics.* The Society of Actuaries. Itasca, Illinois.

[5] Davis, M.H.A. (1993). *Markov Models and Optimization.* Chapman & Hall.

[6] De Vylder, F. and Jaumain, C. (1976). *Exposé moderne de la théorie mathématique des opérations viagère.* Office des Assureurs de Belgique, Bruxelles.

[7] Föllmer, H., Sondermann, D. (1986): Hedging of non-redundant claims. In *Contributions to Mathematical Economics in Honor of Gerard Debreu*, 205-223, eds. Hildebrand, W., Mas-Collel, A., North-Holland.

[8] Gerber, H.U. (1995). *Life Insurance Mathematics*, 2nd edn. Springer-Verlag.

[9] Hald, A. (1987). On the early history of life insurance mathematics. *Scandinavian Actuarial Journal* **1987**, 4-18.

[10] Hoem, J.M. (1969): Markov chain models in life insurance. *Blätter der Deutschen Gesellschaft für Versicherungsmathematik* **9**, 91–107.

[11] Hoem, J.M. and Aalen, O.O. (1978). Actuarial values of payment streams. *Scandinavian Actuarial Journal* **1978**, 38-47.

[12] Jordan, C.W. (1967). *Life Contingencies.* The Society of Actuaries, Chicago.

[13] Møller, T. (1998): Risk minimizing hedging strategies for unit-linked life insurance. *ASTIN Bulletin* **28**, 17-47.

[14] Norberg, R. (1991). Reserves in life and pension insurance. *Scandinavian Actuarial Journal* **1991**, 1-22.

[15] Norberg, R. (1999): A theory of bonus in life insurance. *Finance and Stochastics.* **3**, 373-390.

[16] Steffensen, M. (2000): A no arbitrage approach to Thiele's differential equation. *Insurance: Mathematics & Economics* **27**, 201-214.

Ragnar Norberg

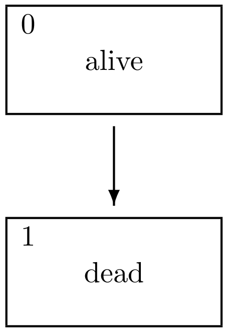


Figure 1: A single-life policy with two states.

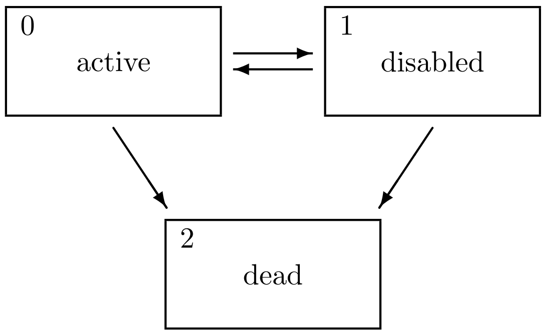


Figure 2: A single-life policy with three states.

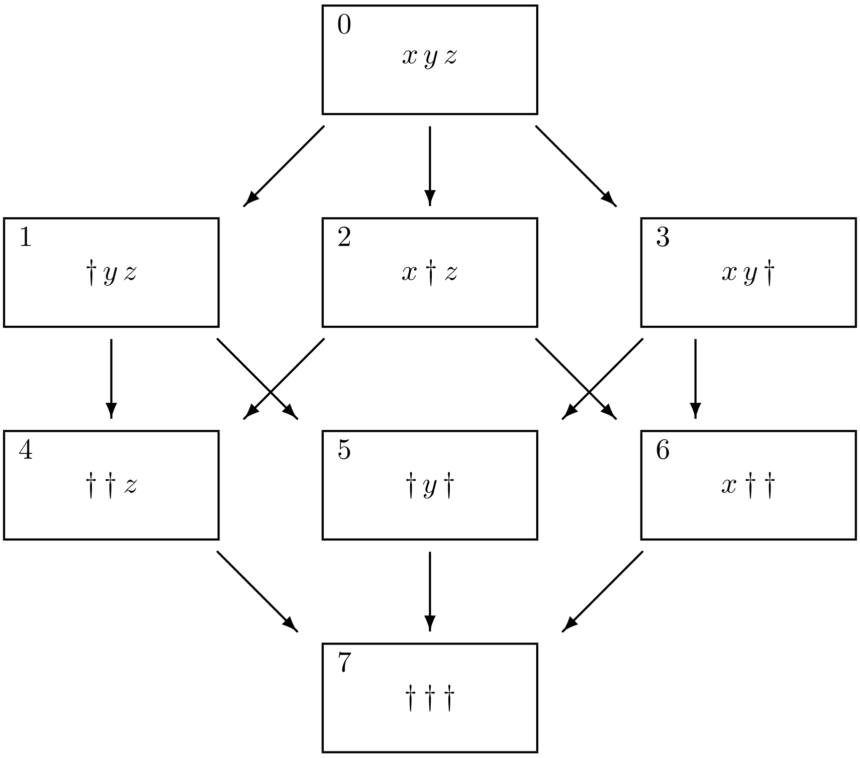


Figure 3: A policy involving three lives $x$, $y$, $z$. An expired life is replaced by a dagger †.