Chapter 3
Shocks and Degradation

This chapter is mostly devoted to basic shock models and their simplest applications. Along with discussing some general approaches and results, we want to present the necessary material for describing our recent findings on shocks modeling of the next chapter. As in the other chapters of this book, we do not intend to perform a comprehensive literature review of this topic, but rather concentrate on notions and results that are vital for further presentation.

We understand the term "shock" in a very broad sense as some instantaneous, potentially harmful event (e.g., electrical impulses of large magnitude, demands for energy in biological objects, insurance claims in finance, etc.). Shock models are widely used in practical and theoretical reliability and in the other disciplines as well. They can also constitute a useful framework for studying aging properties of distributions [2, 3]. It is important to analyze the consequences of shocks to a system (object) that can be *basically* two fold. First, under certain assumptions, we can consider shocks that can either 'kill' a system, or be successfully survived without any impact on its future performance. The corresponding models are usually called the *extreme shock models*, whereas the setting when each shock results in an additive damage (wear) to a system is often described in terms of the *cumulative shock models* ([18–20] to name a few). In the latter case, the failure occurs when the cumulative effect of shocks reaches some deterministic or random level, and therefore, this setting is useful for modeling of degradation (wear) processes. The combination of these two basic models has been also considered in the literature [5, 6, 19].

In Sect. 3.1, we first briefly discuss several simplest stochastic models of wear that are helpful in describing basic cumulative shock models. In the rest of this chapter, we mostly consider the basic results with respect to the extreme and cumulative shock models, and also describe several meaningful modifications, and applications of the extreme shock model. For instance, in Sect. 3.8, a meaningful safety at sea application is considered and in Sect. 3.9, the famous in demography Strehler–Mildvan model of human mortality is discussed from our view point.

M. Finkelstein and J. H. Cha, *Stochastic Modeling for Reliability*,
Springer Series in Reliability Engineering, DOI: 10.1007/978-1-4471-5028-2_3,

3.1 Degradation as Stochastic Process

Stochastic degradation in engineering, ecological, and biological systems is naturally modeled by increasing (decreasing) stochastic processes. The additive nature of the cumulative shock models implies that the corresponding degradation should be strictly monotone. However, it is well-known (e.g., [3] that, for example, the Wiener process with drift (see Definition 3.1) with the nonmonotone realizations under certain assumptions can be also considered as a useful tool for modeling the monotone degradation. In the previous chapter, several point processes were discussed that can be used for modeling degradation induced by shocks in the corresponding cumulative shock models. We will consider now the simplest *continuous-time* stochastic processes, and will be interested in modeling stochastic degradation as such and in obtaining the corresponding distributions for the first passage times when this degradation reaches the predetermined or random level D for the first time. When D defines some critical safety boundary, the latter interpretation can be useful for risk and safety assessment. For instance, when degradation in some structures results in the decreasing resistance to loads, it can result not just in an 'ordinary' failure, but in a severe catastrophic event.

We will briefly define now several approaches, which are most often used in engineering practice for degradation modeling. The simplest and the widely used one is the path model. Its stochastic nature is described either by the additive or by the multiplicative random variable in the following way:

$$W_t = \eta(t) + Z, \tag{3.1}$$

$$W_t = \eta(t)Z, \tag{3.2}$$

where $\{W_t,\ t \geq 0\}$ denotes our stochastic process, $\eta(t)$ is an increasing, continuous function ($\eta(0) = 0,\ \lim_{t\to\infty} \eta(t) = \infty$) and Z is a nonnegative random variable with the Cdf $G(z)$. Therefore, the sample paths (realizations) for these models are monotonically increasing. The 'nature' of this stochastic process is simple and meaningful: let the failure (catastrophe) be defined as reaching by $\{W_t, t \geq 0\}$ the degradation threshold $D > 0$ and T_D be the corresponding time to failure random variable with the Cdf $F_D(t)$. It follows, e.g., for the model (3.2) that:

$$F_D(t) = P(W_t \geq D) = \Pr\left(Z \geq \frac{D}{\eta(t)}\right) = 1 - G\left(\frac{D}{\eta(t)}\right). \tag{3.3}$$

Example 3.1 Let $\eta(t) = t$ and assume that Z is described by the Weibull distribution, i.e., $G(z) = 1 - \exp\left\{-(\lambda z)^k\right\}$, $\lambda, k > 0$. Then, in accordance with (3.3),

$$F_D(t) = \exp\left\{-\left(\frac{\lambda D}{t}\right)^k\right\},$$

which is often called the Inverse-Weibull distribution [1]. Specifically, when $\lambda = 1, k = 1$:

$$F_D(t) = \exp\left\{-\frac{D}{t}\right\}.$$

It is clear that the value at $t = 0$ for this distribution should be understood as

$$F_D(0) = \lim_{t\to 0} F_D(t) = 0.$$

The Inverse-Weibull distribution is a convenient simple tool for describing threshold models with a linear function $\eta(t)$.

Assume now that the threshold D is a random variable with the Cdf $F_0(d) = \Pr(D \le d)$ and let, at first, degradation be modeled by the deterministic, increasing function $W(t)$ ($W(0) = 0$, $\lim_{t\to\infty} W(t) = \infty$). Equivalently, the problem can be reformulated in terms of the fixed threshold and random initial value of degradation. Denote by T the random time to failure. As events $T \le t$ and $W(t)$ are equivalent, similar to (3.3) [12],

$$F(t) \equiv P(T \le t) = P(D \le W(t)) = F_0(W(t)), \tag{3.4}$$

where the last equality is due to the fact that the Cdf of D is $F_0(d)$. Substituting d by $W(t)$, finally results in (3.4).

Let now the deterministic degradation $W(t)$ in (3.4) be replaced by a stochastic process $W_t, t \ge 0$. In order to derive the corresponding distribution of the time to failure in this case we must obtain the expectation of $F_0(W_t)$ with respect to the process $W_t, t \ge 0$:

$$F(t) = E\left[F_0(W)_t\right]. \tag{3.5}$$

This equation is too general, as the stochastic process is not specified. The following example considers the multiplicative path model for $W_t, t \ge 0$.

Example 3.2 Let, e.g., $F_0(d) = 1 - \exp\{-\lambda d\}$ and $W_t = \eta(t)Z$, where Z is also exponentially distributed with parameter μ. Direct integration in (3.5) gives:

$$\begin{aligned} F(t) &= E[1 - \exp\{-\lambda\eta(t)Z\}] \\ &= \int_0^\infty (1 - \exp\{-\lambda\eta(t)z\})\mu\exp\{-\mu z\} \\ &= 1 - \frac{\mu}{\mu + \lambda\eta(t)}. \end{aligned}$$

The path model can be very useful for illustration. However, obviously, the real life stochastic processes are much more complex. Probably, the most popular in applications and well investigated from the formal point of view stochastic process is the Wiener process. The Wiener process with drift is often used for modeling wear although its sample paths are not monotone (but the mean of the process is a monotonically increasing function).

Definition 3.1 Stochastic process $\{W_t, t \geq 0\}$ is called the Wiener process with drift

$$W_t = \mu t + X(t),$$

where $\mu > 0$ is a drift parameter and $X(t)$ is a standard Wiener process: for the fixed $t \geq 0$, the random variable $X(t)$ is normally distributed with zero mean and variance $\sigma^2 t$.

It is well-known (see, e.g., Cox and Miller [8] that the first passage time T_D, i.e.,

$$T_D = \inf_t \{t, W_t > D\}$$

for this process is described by the inverse Gaussian distribution:

$$\overline{F}_D(t) = \Pr(T_D > t) = \Phi\left(\frac{D - \mu t}{\sqrt{t}\,\sigma}\right) - \exp\{-2D\mu\}\Phi\left(\frac{D + \mu t}{\sqrt{t}\,\sigma}\right) \tag{3.6}$$

and

$$E[T_D] = \frac{D}{\mu}, \quad Var(T_D) = \frac{D\sigma^2}{\mu^3},$$

where, as usual, $\Phi(t)$, denotes the Cdf of the standard normal random variable.

Another popular process for modeling degradation is the gamma process (see, e.g., the perfect survey by Van Nortwijk [30]). Although, parameter estimation for the degradation models driven by the gamma process is usually more complicated than for the Wiener process, it better captures the desired monotonicity.

Definition 3.2 The gamma process is a stochastic process $(W_t, t \geq 0)$, $W_0 = 0$ with independent nonnegative increments having a gamma Cdf with identical scale parameters. The increment $W_t - W_\tau$ has a gamma distribution with a shape parameter $v(t) - v(\tau)$ and a scale parameter u, where $v(t)$ is an increasing function ($v(0) = 0$).

Thus W_t for each fixed t is gamma-distributed with shape parameter $v(t)$ and scale parameter u, whereas

$$E[W_t] = \frac{v(t)}{u}, \quad Var(W_t) = \frac{v(t)}{u^2}.$$

The first passage time T_D, is described in this case by the following distribution [30]

$$F_D(t) = \Pr(T_D \le t) = \Pr(W_t \ge D) = \frac{\Gamma(v(t), Du)}{\Gamma(v(t))},$$

where $\Gamma(a,x) = \int_x^\infty t^{a-1}e^{-t}\mathrm{d}t$ is an incomplete gamma function for $x > 0$. Thus, deterioration with independent increments can be often modeled by the gamma process.

3.2 Shocks and Shot Noise Process

A natural way of modeling additive degradation is via the sum of random variables, which represent the degradation increments:

$$W_t = \sum_1^n X_i,$$

where $X_i, i = 1, 2, \ldots, n$ are positive i.i.d. random variables with a generic variable denoted by X, and n is an integer.

The next step to a more real stochastic modeling is to view n as a random variable N (the compound random variable) or a point process $\{N_t, t \ge 0\}$. The latter is counting the point events of interest in $[0, t), t \ge 0$ (the compound point process):

$$W_t = \sum_1^{N_t} X_i. \tag{3.7}$$

Denote by $Y_i, i = 1, 2, \ldots$ a sequence of inter-arrival times for $\{N_t, t \ge 0\}$. If $Y_i, i = 1, 2, \ldots$ are i.i.d (and this case will be considered in what follows) with a generic variable Y, then the Wald's equation [26] immediately yields

$$E[W_t] = E[N_t]E[X],$$

where, specifically for the compound Poisson process with rate m: $E[N_t] = mt$. Note that [9] under certain assumptions the stationary gamma process ($v(t) = vt$) can be viewed as a limit of a specially constructed compound Poisson process.

Relationship (3.7) has a meaningful interpretation via shocks, as $X_i, i = 1, 2, \ldots$ can be interpreted as an amount of damage caused by the ith shock. An important modification of this additive model is given by the shot noise process [25, 26]. In a shot noise point process, an additive input of a shock of magnitude X_i is decreased in accordance with some decreasing (nonincreasing) response function $h(t - s)$. Therefore, Eq. (3.7) turns to

$$W_t = \sum_1^{N_t} X_i h(t - \tau_i), \tag{3.8}$$

where $\tau_1 < \tau_2 < \tau_3, \ldots$ is the sequence of the corresponding arrival (waiting) times in the point process. This setting has a lot of applications in electrical engineering, materials science, health sciences, risk, and safety analysis. For instance, cracks due to fatigue in some materials tend to close up after the material has borne a load, which has caused the cracks to grow. Another example is the human heart muscle's tendency to heal after a heart attack [27]. Thus, the inputs of each shock in the accumulated damage decrease with time.

Equivalently, (3.8) can be written as:

$$W_t = \int_0^t Xh(t-u)dN_u,$$

where $dN_u = N(u, u + \mathrm{d}u)$ denotes the number of shocks in $[u, u + \mathrm{d}u)$.

First, we are interested in the mean of the defined process. Assume that $E[X] < \infty$. As $X_i, i = 1, 2, \ldots$ are independent from the point process $\{N_t, t \geq 0\}$,

$$E[W_t] = E[X] \int_0^t h(t-u)dN_u = E[X] \int_0^t h(t-u)m(u)\mathrm{d}u, \tag{3.9}$$

where $m(u) = dE[N_u]/\mathrm{d}u$ is the rate (intensity) of the point process. For the Poisson process, $m(u) = m$ and:

$$E[W_t] = mE[X] \int_0^t h(u)\mathrm{d}u. \tag{3.10}$$

Therefore, asymptotically the mean accumulative damage is finite, when the response function has a finite integral, i.e.,

$$\lim_{t \to \infty} E[W_t] < \infty \ , \ if \quad \int_0^\infty h(u)\mathrm{d}u < \infty.$$

This property has an important meaning in different engineering and biological applications. It can be shown directly that, if $E[X^2] < \infty$:

$$Cov(W_{t_1}, W_{t_2}) = mE\left[X^2\right] \int_0^{t_1} h(t_1 - u)h(t_2 - u)\mathrm{d}u; \quad t_1 \geq t_2.$$

The central limit theorem for the sufficiently large m also takes place in the following form [23, 24]:

$$\frac{W_t - E[W_t]}{(Var(W_t))^{1/2}} \to^D \ N(0,1), \ t \to \infty \ , \tag{3.11}$$

where the sign "D" means convergence in distribution and $N(0,1)$ denotes the standard normal distribution. The renewal case with the interarrival time denoted by X gives similar results

$$\lim_{t\to\infty} E[W_t] = \frac{1}{E[X]}\int_0^\infty h(u)\mathrm{d}u.$$

Example 3.3 Consider a specific exponential case of the response function $h(u)$ and the Poisson process of shocks with rate m:

$$W_t = \sum_1^{N_t} X_i \, \exp\{\alpha(t-\tau_i)\}.$$

By straightforward calculations [26], using the technique of the moment generating functions, it can be shown that the stationary value of W_t for t sufficiently large is described by the gamma distribution with mean $m/\lambda\alpha$ and variance $m/\lambda^2\alpha$. Moreover, the distribution of the first passage time is given by

$$F_D(t) = \Pr(T_D \le t) = \Pr(W_t \ge D) = \frac{\Gamma(m/\alpha, D\lambda)}{\Gamma(m/\alpha)}.$$

It is well-known from the properties of the gamma distribution that as m/λ increases, it converges to the normal distribution and, therefore, there is no contradiction between this result and asymptotic relation (3.11).

In the next chapter, we will consider another shot noise model where the shot-noise process models the failure rate of an object. Some meaningful generalizations will be also considered.

3.3 Asymptotic Properties

In many applications, the number of shocks in the time interval of interest is large, which makes it possible to apply the corresponding asymptotic methods.

Consider a family of nonnegative, i.i.d, two-dimensional random vectors $\{(X_i, Y_i), i \ge 0\}, X_0 = 0, Y_0 = 0$, where $\sum_1^n X_i$ is the accumulated damage after n shocks and $Y_i, i = 1,2,\ldots$ is the sequence of the i.i.d inter-arrival times of the corresponding renewal process. Recall that the renewal process is defined by the sequence of the i.i.d inter-arrival times. Specifically, when these times are exponentially distributed, the renewal process 'reduces' to the Poisson process. We will assume for simplicity that X and Y are independent, although the case of dependent variables can be also considered [19]. Let $0 < E[X],\ E[Y] < \infty,\ 0 < Var(X),\ \ Var(Y) < \infty$. It follows immediately from (3.7) and the elementary renewal theorem [26] that

$$\lim_{t\to\infty}\frac{E[W_t]}{t}=\lim_{t\to\infty}\frac{E[N_t]E[X]}{t}=\frac{E[X]}{E[Y]}. \tag{3.12}$$

The corresponding central limit theorem can be proved using the theory of stopped random walks [19]

$$\frac{W_t-(E[X]/E[Y])t}{(E[Y])^{-3/2}\sigma\, t^{1/2}}\to\ N(0,1),\ t\to\infty. \tag{3.13}$$

where $\sigma=\sqrt{\mathrm{var}(E[Y]X-E[X]Y)}$.

Relationship (3.13) means that for large t, the random variable W_t is approximately normally distributed with expected value $(E[X]/E[Y])t$ and variance $(E[Y])^{-3}\ \sigma^2(E[X])^2\,t$. Therefore, we need only $E[X]$, $E[Y]$ and σ for the corresponding asymptotic analysis, which is very convenient in practice.

Similar to (3.12),

$$\lim_{t\to\infty}\frac{E[T_D]}{D}=\lim_{D\to\infty}\frac{E[N_D]E[Y]}{D}=\frac{E[Y]}{E[X]}, \tag{3.14}$$

where N_D denotes a random number of shocks to reach the cumulative value D. Equation (3.13) can be now rewritten for the distribution of the first passage time T_D as [19]

$$\frac{T_D-(E[Y]/E[X])D}{(E[X])^{-3/2}\,\sigma\,D^{1/2}}\to\ N(0,1),\ D\to\infty.$$

This equation means that for large threshold D the random variable T_D can be approximately described by a normal distribution with expected value $(E[Y]/E[X])D$, and variance $(E[X])^{-3}\ \sigma^2\ D$. Therefore, the results of this section can be easily and effectively used in safety and reliability analysis.

3.4 Extreme Shock Models

Let the shocks occur in accordance with a renewal process or a nonhomogeneous Poisson process. Each shock *independently of the previous history* leads to a failure of a system with probability p and is survived with the complementary probability $q=1-p$. Assume, that a shock is the only cause of failure. We see that there is no accumulation of damage and the fatal 'damage' can be a consequence of a single shock. Numerous problems in reliability, risk, and safety analysis can be interpreted by means of this model. This setting is often referred to as an *extreme shock model* [12, 18]. Our main interest in the rest of this chapter will be in different settings, and applications that are described within the framework of the extreme shock model. We will use these results and reasoning in the rest of this book.

Consider first, a general point process $\{T_n\}; T_0 = 0, T_{n+1} > T_n, n = 0, 1, 2, \ldots$, where T_n is the time to the nth arrival of an event with the corresponding cumulative distribution function $F^{(n)}(t)$. Therefore, $F^{(n)}(t) - F^{(n+1)}(t)$ is the probability of exactly n events in $[0, t)$; $F^{(0)}(t) \equiv 1$, $F^{(1)}(t) \equiv F(t)$. Let G be a geometric variable with parameter p (independent of $\{T_n\}_{n \geq 0}$) and denote by T a random variable with the following survival function

$$P(t) = \sum_{k=0}^{\infty} q^k \left(F^{(k)}(t) - F^{(k+1)}(t) \right). \tag{3.15}$$

Thus $P(t)$ is the system's survival probability for the described extreme shock model. We can also interpret the setting in terms of the terminating point process when $1 - P(t)$ is the probability of its termination in $[0, t)$.

Obtaining probability $P(t)$ is an important problem in various reliability and safety assessment applications. It is clear that in this general form, Eq. (3.15) does not allow for explicit results that can be used in practice, and therefore, assumptions on the type of the point process of shocks should be made. Two specific point processes are mostly used in reliability applications, i.e., the Poisson process and the renewal process. For the homogeneous Poisson process with rate λ, the derivation is trivial

$$P(t) = \sum_{0}^{\infty} q^k \ \exp\{-\lambda t\} \frac{(\lambda t)^k}{k!} = \exp\{-p\lambda t\}. \tag{3.16}$$

It follows from (3.16) that the corresponding constant failure rate, which describes the lifetime of our system T, is given by a simple and meaningful relationship

$$\lambda_S = p\,\lambda. \tag{3.17}$$

Thus, the rate of the underlying Poisson process λ is decreased by the factor $p \leq 1$.

This result can be generalized to the case of the NHPP with rate $\lambda(t)$ and time-dependent probability $p(t)$. It is clear that the Brown–Proschan model of Chap. 2 described by Eqs. (2.17–2.19) can be interpreted in terms of our extreme shock model, and therefore,

$$P(t) = 1 - \exp\left\{ - \int_0^t p(u)\lambda(u)\mathrm{d}u \right\} \tag{3.18}$$

with the corresponding failure rate

$$\lambda_S(t) = p(t)\lambda(t).$$

Numerous generalizations of these results under the assumption of the underlying NHPP of shocks will be considered further in this chapter and in the next chapter as well. In spite of its relative simplicity, the renewal process of shocks does not allow for the similar explicit relationships. However, it is well-known (see, e.g., [21]) that, as $p \to 0$, the following convergence in distribution takes place:

$$P(t) \to \exp\left\{-\frac{pt}{\mu}\right\}, \quad \forall t \in (0, \infty) \ , \tag{3.19}$$

where μ is the mean that corresponds to the governing distribution. Thus, (3.19) constitutes a very simple asymptotic exponential approximation. In practice, however, parameter p is not usually sufficiently small for using effectively this approximation, and therefore, the corresponding bounds for $P(t)$ can be very helpful.

The simplest and useful in practice but a rather crude bound for the survival function can be obtained via the following identity:

$$E\left[q^{N_t}\right] = \sum_{k=0}^{\infty} q^k \left(F^{(k)}(t) - F^{(k+1)}(t)\right).$$

Finally, using Jensen's inequality [12]:

$$P(t) = E\left[q^{N_t}\right] \geq q^{E[N_t]}.$$

In the next three sections, the extreme shock model with the *homogeneous* Poisson process of shocks will be generalized to different settings that can occur in practice [13]. For instance, the probability of a failure of an operable system under a shock, which is in conventional models either a constant or depends only on chronological time t, can depend also on a state of a system. This is a natural assumption, as resistance to shocks, e.g., in multistate systems (discrete or continuous) often depends on the current state of a system. Another extension of conventional models to be considered is when the failure occurs if two successive shocks 'are too close' to each other. A system in this case cannot recover from the consequences of the previous shock. This setting is similar to that of the σ -shock model considered in the literature [22, 28], however, our method allows for more general and flexible results. The main analytical tool allowing for the *explicit solutions* for all mentioned settings is the method of integral equations developed in Finkelstein [12]. These equations can be effectively solved in terms of the Laplace transform and explicitly inverted for the sufficiently simple cases.

3.5 State-Dependent Probability of Termination

Consider first, the Poisson process of shocks with rate λ and probability of failure (termination) on each shock, p. In this case, the survival probability is given by Eq. (3.16). In order, to illustrate the *method of integral equations* to be used further [13] we will describe how it works for this simplest case. It is easy to see that the following integral equation with respect to $P(t)$ holds

$$P(t) = e^{-\lambda t} + \int_0^t \lambda e^{-\lambda x} q\, P(t-x)\mathrm{d}x \,. \tag{3.20}$$

The first term, on the right hand side is the probability that there are no shocks in $[0, t)$ and the integrand defines the probability that the first shock that have occurred in $[x, x + \mathrm{d}x)$ was survived and then the system have survived in $[x, t)$. Due to the properties of the homogeneous Poisson process, the probability of the latter event is $P(t-x)$.

We have now a simple integral equation with respect to the unknown function $P(t)$. Applying the Laplace transform to both sides of Eq. (3.20) results in

$$\widetilde{P}(s) = \frac{1}{s + \lambda} + \frac{\lambda q}{s + \lambda}\widetilde{P}(s) \Rightarrow \widetilde{P}(s) = \frac{1}{s + \lambda p},$$

where $\widetilde{P}(s)$ denotes the Laplace transform of $P(t)$. The corresponding inversion results in $\exp\{-p\lambda t\}$.

Consider now a repairable system with instantaneous, perfect repair that starts functioning at $t = 0$. Let its lifetime be described by the Cdf $F(t)$, which is a governing distribution for the corresponding renewal process with the renewal density function to be denoted by $h(t)$. Assume, that the quality of performance of our system is characterized by some deterministic for simplicity function of performance $Q(t)$ to be called the quality function. The considered approach can be generalized to the case of a random $Q(t)$. It is often a decreasing function of time, and this assumption is quite natural for degrading systems. In applications, the function $Q(t)$ can describe some key parameter of a system, e.g., the decreasing in time accuracy of the information measuring system or effectiveness (productivity) of some production process. As repair is perfect, the quality function is also restored to its initial value $Q(0)$. It is clear that the quality function of our system at time t is now random and equal to $Q(Y)$, where Y is a random time since the last (before t) repair.

The system is subject to the Poisson process of shocks with rate λ. As previously, each shock can terminate the performance of the repairable system and we are interested in obtaining the survival probability $P(t)$. Note, that the repaired failure of the system does not terminate the process and only a shock can result in termination. Assume, that the probability of termination depends on the system's quality at the time of a shock. This is a reasonable assumption meaning that the larger value of quality implies the smaller probability of termination. Let the first shock arrive before the first failure of the system. Denote by $p^*(Q(t))$ the corresponding probability of termination in this case. Now we are able to obtain $p(t)$—the probability of termination of the operating system *by the first shock* at time instant t. Using the standard 'renewal-type reasoning' [13], the following relationship for $p(t)$ can be derived

$$p(t) = p^*(Q(t))\,\overline{F}(t) + \int_0^t h(x)\,\overline{F}(t-x)\,p^*(Q(t-x))\mathrm{d}x \,, \tag{3.21}$$

where $\overline{F}(t) \equiv 1 - F(t)$.

The first term on the right-hand side of Eq. (3.21) gives the probability of termination during the first cycle of the renewal process, whereas $h(x)\,\overline{F}\,(t-x)\mathrm{d}x$ defines the probability that the last failure (renewal) of the system before t had occurred in $[x, x + \mathrm{d}x)$ (as $h(x)\mathrm{d}x$ is the probability that a failure (renewal) had occurred in $[x, x + \mathrm{d}x)$ and $\overline{F}(t-x)$ is the probability that no failure had occurred in $[x + \mathrm{d}x, t]$. Therefore, the corresponding probability of termination at t is equal to $p^*(Q(t-x))$.

Thus, the probability of termination under the first shock $p(t)$, which is now time-dependent, has been derived. Assume, now that the survived shock can be interpreted as an instantaneous, perfect repair of the system (the 'repaired shock' is survived, the 'non-repaired' results in termination). Therefore, the instants of survived shocks can be also considered as the renewal points for the system. Having this in mind, we can now proceed with obtaining the survival probability $P(t)$. Using the similar reasoning as when deriving Eq. (3.20)

$$P(t) \; = \; e^{-\lambda t} \; + \int_0^t \lambda e^{-\lambda x} \, q(x) \, P(t-x)\mathrm{d}x \; , \tag{3.22}$$

where $q(x) \equiv 1 - p(x)$.

Applying the Laplace transform to Eq. (3.22):

$$\begin{aligned} \widetilde{P}(s) &= \frac{1}{s+\lambda} + \lambda \widetilde{q}(s+\lambda)\widetilde{P}(s) \\ \Rightarrow \widetilde{P}(s) &= \frac{1}{(s+\lambda)(1-\lambda\widetilde{q}(s+\lambda))}. \end{aligned} \tag{3.23}$$

Given the functions $F(t)$ and $p^*(Q(t))$, Eqs. (3.21) and (3.23) can be solved numerically, but we can still proceed with the Laplace transforms under an additional assumption that the underlying distribution is exponential, i.e., $F(t) = 1 - \exp\{-ht\}$. In this case, $h(x) = h$ and the Laplace transform of Eq. (3.21) results in [13]

$$\widetilde{p}(s) = \widetilde{p}^*(s+h)\left(1 + \frac{h}{s}\right), \tag{3.24}$$

where $\widetilde{p}^*(s) = \int_0^\infty e^{-sx} p^*(Q(x))\mathrm{d}x$ denotes the Laplace transform of the function $p^*(Q(t))$. Substituting (3.24) into (3.23) and taking into account that $\widetilde{q}(s) = (1/s) - \widetilde{p}(s)$

$$\widetilde{P}(s) = \frac{1}{s + \lambda\,\widetilde{p}^*(s+h+\lambda)(s+h+\lambda)}. \tag{3.25}$$

To proceed further with inversion, we must make some assumptions on the form of the function $p^*(Q(t))$. Let $p^*(Q(t)) = 1 - \exp\{-\alpha t\}$, $\alpha \geq 0$. This is a reasonable assumption (as the probability of termination increases as $Q(t)$ decreases with t) that allows for a simple Laplace transform. Then

$$\widetilde{P}(s) = \frac{s + h + \lambda + \alpha}{s^2 + s(\lambda + h + \alpha) + \alpha\lambda}$$

and the inversion gives

$$P(t) = \frac{s_1 + \lambda + \alpha}{s_1 - s_2} \exp\{s_1 t\} - \frac{s_2 + \lambda + \alpha}{s_1 - s_2} \exp\{s_2 t\},$$

where

$$s_{1,2} = \frac{-(h + \lambda + \alpha) \pm \sqrt{(h + \lambda + \alpha)^2 - 4\lambda\alpha}}{2}.$$

An important specific case is when the system is absolutely reliable ($h = 0$) but is characterized by the quality function $Q(t)$. Then $s_1 = -\lambda$, $s_2 = -\alpha$; $\alpha \neq \lambda$ and

$$P(t) = \frac{\lambda}{\lambda - \alpha} \exp\{-\alpha t\} - \frac{\alpha}{\lambda - \alpha} \exp\{-\lambda t\}. \tag{3.26}$$

If, for instance, $p^*(Q(t)) = 1$, which means that $\alpha \to \infty$, then $P(t) = \exp\{-\lambda t\}$ as expected, the probability that there are no shocks in $[0, t)$. On the contrary, if $\alpha = 0$, which means that $p^*(Q(t)) = 0$, the survival probability is equal to 1. Another marginal case is defined by the value of the rate λ. If $\lambda = 0$, then again, as expected, $P(t) = 1$. On the other hand, it follows from (3.26) that as $\lambda \to \infty$,

$$P(t) \to \exp\{-\alpha t\}, \tag{3.27}$$

which can be confusing at first sight, as one would expect that when the rate of a shock process tends to infinity, the probability of survival in $[0, t)$ should tend to 0, but this is not the case because the function $p^*(Q(t)) = 1 - \exp\{-\alpha t\}$ is close to 0 for small t and each survived shock is the renewal point for our system. Therefore, as the number of shocks increases, due to the properties of exponential function, relationship (3.27) holds.

3.6 Termination with Recovery Time

In the previous sections, the only source of termination was an immediate effect of a shock. Consider now another setting that can be often encountered in practical reliability and safety analysis. Let, as previously, each shock from the Poisson process with rate λ terminate the process with probability p and be survived with probability $q = 1 - p$. Assume, now that termination additionally can also occur when the consecutive shocks are 'too close', which means that the system cannot recover from the consequences of a previous shock. Therefore, the time for recovering should be taken into account. It is natural to assume that it is a random variable τ with the Cdf $R(t)$ (different values of damage need different time of recovering and this fact is described by $R(t)$). Thus, if the shock occurs while the system still has not recovered from the previous non-terminating shock, it terminates the process. It is the simplest criterion of termination of this kind. Other criterions can be also considered. As previously, we want to derive $P(t)$—the probability of survival of our system in $[0, t)$.

First, assume that a shock had occurred at $t = 0$ and has been survived. Denote the probability of survival under this condition by $P^*(t)$. Then the corresponding supplementary integral equation is

$$P^*(t) = e^{-\lambda t} + \int_0^t \lambda e^{-\lambda x} q R(x) P^*(t - x)\mathrm{d}x\ , \tag{3.28}$$

where the multiplier $R(x)$ in the integrand is the probability that the recovery time after the first shock at $t = 0$ (and before the next one at $t = x$) is sufficient (smaller than x).

Applying, the Laplace transform to both sides of (3.28) results in the following relationship for the Laplace transform of $P^*(t)$:

$$\tilde{P}^*(s) = \frac{1}{(s + \lambda)(1 - \lambda q \tilde{R}(s + \lambda))}, \tag{3.29}$$

where $\tilde{R}(s)$ is the Laplace transform of the Cdf $R(t)$.

Using probability $P^*(t)$, we can derive now the following equation:

$$P(t) = e^{-\lambda t} + \int_0^t \lambda e^{-\lambda x} q P^*(t - x)\mathrm{d}x. \tag{3.30}$$

As previously, the first term on the right-hand side of this equation is the probability of shocks absence in $[0, t)$, $\lambda e^{-\lambda x} q \mathrm{d}x$ is the probability that the first shock has occurred and was survived in $[x, x + \mathrm{d}x)$. Finally, $P^*(t - x)$ is the probability that the system survives in $[x, t)$.

We can obtain $P(t)$, applying the Laplace transform to both sides of (3.30), i.e.,

$$\tilde{P}(s) = \frac{1}{s + \lambda} + \frac{\lambda q}{s + \lambda} \tilde{P}^*(s)\ ,$$

where $\tilde{P}^*(s)$ is defined by (3.29). This gives the general solution of the problem under the stated assumptions in terms of the Laplace transforms. In order to be able to invert $\tilde{P}(s)$, assume additionally that the Cdf $R(t)$ is exponential, i.e., $R(t) = 1 - \exp\{-\gamma t\}$, $\gamma > 0$. Performing simple algebraic transformations

$$\tilde{P}(s) = \frac{s + 2\lambda + \gamma - p\lambda}{s^2 + s(\gamma + 2\lambda) + \lambda^2 + \gamma\lambda p}. \tag{3.31}$$

Inversion of (3.31) gives

$$P(t) = \frac{s_1 + \gamma + 2\lambda - p\lambda}{s_1 - s_2}\exp\{s_1 t\} - \frac{s_2 + \gamma + 2\lambda - p\lambda}{s_1 - s_2}\exp\{s_2 t\}\ , \tag{3.32}$$

where

$$s_{1,2} = \frac{-(\gamma + 2\lambda) \pm \sqrt{(\gamma + 2\lambda)^2 - 4(\lambda^2 + \gamma\lambda p)}}{2}.$$

Equation (3.32) presents the exact solution for $P(t)$. In applications, it is convenient to use simple approximate formulas. Consider the following meaningful assumption [13]:

$$\frac{1}{\lambda} \gg \bar{\tau} \equiv \int_0^\infty (1 - R(x))\mathrm{d}x\ , \tag{3.33}$$

where $\bar{\tau}$ denotes the mean time of recovery.

Relationship (3.33) means that the mean inter-arrival time in the shock process is much larger than the mean time of recovery, and this is often the case in practice. In the study of repairable systems, the similar case is usually called the *fast repair* condition. Using this assumption, the equivalent rate of termination for our process for $\lambda\bar{\tau} \to 0$, $\lambda t \gg 1$ can be written as

$$\lambda(t) = B\,\lambda(1 + o(1)), \tag{3.34}$$

where B is the probability of termination for the *occurred shock* due to two causes, i.e., the termination immediately after the shock and the termination when the next shock occurs before the recovery is completed. Therefore, for sufficiently large t $(t \gg \bar{\tau})$ the integration in the following integral can be performed to ∞ and the approximate value of B is

$$B = \theta + (1 - \theta)\int_0^\infty \lambda e^{-\lambda x}(1 - R(x))\mathrm{d}x.$$

Assuming, as previously, that $R(t) = 1 - \exp\{-\gamma t\},\ \gamma > 0$ gives

$$B = \frac{\lambda + \theta\gamma}{\lambda + \gamma}.$$

Finally, the fast repair approximation for the survival probability is

$$P(t) \approx \exp\left\{-\frac{\lambda + p\gamma}{\lambda + \gamma}\lambda t\right\}. \tag{3.35}$$

It can be easily seen that when $\gamma \to \infty$ (instant recovery), Relationship (3.35) reduces to Eq. (3.16). The accuracy of the fast repair approximation (3.35) with respect to the time of recovery can be analyzed similar to Finkelstein and Zarudnij [14].

3.7 Two Types of Shocks

Assume now that there are two types of shocks [13]. As in the previous section, potentially harmful shocks (to be called *redshocks*) result in termination of the process when they are 'too close', i.e., when the time between two consecutive red shocks is smaller than a recovery time with the Cdf $R(t)$. Therefore, in this case, the system does not have enough time to recover from the consequences of the previous red shock. Assume for simplicity that the probability of immediate termination on red shock's occurrence is equal to 0 ($p = 0$). The model can be easily generalized to the case when $p \neq 0$. On the other hand, our system is subject to the process of 'good' (*blue*) shocks. If the blue shock follows the red shock, termination cannot happen no matter how soon the next red shock will occur. Therefore, the blue shock can be considered as a kind of an additional recovery action.

Denote by λ and β the rates of the independent Poisson processes of red and blue shocks, respectively. First, assume that the first red shock has already occurred at $t = 0$. An integral equation for the probability of survival in $[0, t)$, $P^*(t)$ for this case is as follows:

$$\begin{aligned} P^*(t) = e^{-\lambda t} + \int_0^t \beta e^{-\beta x} e^{-\lambda x} \int_0^{t-x} \lambda e^{-\lambda y} P^*(t-x-y) \mathrm{d}y \mathrm{d}x \\ + \int_0^t e^{-\beta x} \lambda e^{-\lambda x} R(x) P^*(t-x) \mathrm{d}x, \end{aligned} \tag{3.36}$$

where

- The first term on the right-hand side is the probability that there are no other red shocks in $[0, t)$;
- $\beta e^{-\beta x} e^{-\lambda x} \mathrm{d}x$ is the probability that a blue shock occurs in $[x, x + \mathrm{d}x)$ and no red shocks occur in $(0, x)$;
- $\lambda e^{-\lambda y} \mathrm{d}y$ is the probability that the second red shock occurs in $[x + y, x + y + \mathrm{d}y)$;
- $P^*(t - x - y)$ is the probability that the system survives in $[x + y, t)$ given the red shock has occurred at time $x + y$;
- $e^{-\beta x} \lambda e^{-\lambda x} \mathrm{d}x$ is the probability that there is one red shock (the second) in $(0, t)$ and no blue shocks in this interval of time;

- $R(x)$ is the probability that the recovery time x is sufficient and, therefore, the second red shock does not terminate the process;
- $P^*(t-x)$ is the probability that the system survives in $[x, t)$ given the red shock has occurred at time x.

Using $P^*(t)$ that can be obtained from Eq. (3.36), as previously, we can now construct an equation with respect to $P(t)$—the probability of survival without assuming occurrence of the red shock at $t = 0$. Thus

$$P(t) = e^{-\lambda t} + \int_0^t \lambda e^{-\lambda x} P^*(t-x)\mathrm{d}x\,. \tag{3.37}$$

Applying the Laplace transform to Eq. (3.36) results in

$$\widetilde{P}^*(s) = \frac{s+\beta+\lambda}{(s+\beta+\lambda)(s+\lambda)-\beta\lambda-\lambda(s+\beta+\lambda)(s+\lambda)\widetilde{R}(s+\beta+\lambda)}. \tag{3.38}$$

Applying the Laplace transform to Eq. (3.38) gives

$$\widetilde{P}(s) = \frac{1}{s+\lambda} + \frac{\lambda}{s+\lambda}\widetilde{P}^*(s).$$

This equation gives a general solution of the problem under the stated assumptions in terms of the Laplace transforms. In order to be able to invert $\widetilde{P}(s)$, as in the previous section, assume that the Cdf $R(t)$ is exponential $R(t) = 1 - \exp\{-\gamma t\},\ \gamma > 0$. Performing simple algebraic transformations

$$\widetilde{P}(s) = \frac{s+\gamma+\beta+2\lambda}{s^2 + s(\gamma+\beta+2\lambda)+\lambda^2}. \tag{3.39}$$

Inversion of (3.39) results in

$$P(t) = \frac{s_1+\gamma+\beta+2\lambda}{s_1-s_2}\exp\{s_1 t\} - \frac{s_2+\gamma+\beta+2\lambda}{s_1-s_2}\exp\{s_2 t\}\,, \tag{3.40}$$

where

$$s_{1,2} = \frac{-(\gamma+2\lambda+\beta) \pm \sqrt{(\gamma+\beta)^2 + 4\lambda(\gamma+\beta)}}{2}.$$

When $\gamma = 0$, there is no recovery time and the process is terminated when two consecutive red shocks occur.

Equation (3.40) gives an exact solution for $P(t)$. Similar to the previous section, it can be simplified under certain assumptions. Assume that the fast repair condition (3.33) holds. The first red shock cannot terminate the process. The probability that the subsequent shock can result in termination is

$$B = \int_0^t \lambda e^{-\lambda x} \int_0^{t-x} \lambda e^{-\lambda y} e^{-\beta y}(1 - R(y))\mathrm{d}y\mathrm{d}x.$$

For the exponentially distributed time of recovery

$$B = \frac{\lambda}{\lambda + \beta + \gamma} - \frac{\lambda}{\beta + \gamma} e^{-\lambda t} + \frac{\lambda^2}{(\lambda + \beta + \gamma)(\beta + \gamma)} e^{-(\lambda+\beta+\gamma)t}.$$

For the sufficiently large t, $B \approx \lambda/\lambda + \beta + \gamma$ and this approximate value can be used for subsequent shocks as well. Therefore, the relationship

$$P(t) \approx \exp\left\{ -\frac{\lambda^2}{\lambda + \beta + \gamma} t \right\}.$$

is the fast repair approximation in this case.

The considered in Sects. 3.5–3.7 method of integral equations, which is applied to deriving the survival probability for different shock models is an effective tool for obtaining probabilities of interest in situations where the object under consideration has renewal points. As the considered process of shocks is the homogeneous Poisson process, each shock (under some additional assumptions) constitutes these renewal points. When a shock process is the NHPP, there are no renewal points, but the integral equations usually can also be derived. For the illustration, consider the corresponding generalization of Eq. (3.20). Denote by $P(t - x, x)$ the survival probability in $[x, t)$, $x < t$ for the 'remaining shock process' that has started at $t = 0$ and was not terminated by the first shock at time x. Note that this probability depends now not only on $x - t$ as in the homogeneous case, but on x as well. Equation (3.20) is modified now to

$$P(t) = \exp\left\{ -\int_0^t \lambda(u)\mathrm{d}u \right\} + \int_0^t \lambda(x) \exp\left\{ -\int_0^x \lambda(u)\mathrm{d}u \right\} qP(t - x, x)\mathrm{d}x.$$

It can be seen by substitution that

$$P(t - x, x) = \exp\left\{ -p \int_x^t \lambda(u)\mathrm{d}u \right\}, \ 0 \leq x, t$$

is the solution to this equation.

One can formally derive integral equations for other models (with the NHPP process of shocks) considered in this section, however, the corresponding solutions can be obtained only numerically, as the explicit inversions of the Laplace transforms are not possible in these cases.

The method of integral equations can be also obviously applied to the renewal process of shocks, as in this case we also have 'pure renewal points'. For instance, the simplest Eq. (3.20) turns into

$$P(t) = (1 - F(t)) + \int_0^t f(x)\, qP(t - x)\mathrm{d}x\,,$$

where $F(t)$ and $f(t)$ are the Cdf and the pdf of the inter-arrival times, respectively. Applying the Laplace transform gives

$$\widetilde{P}(s) = \frac{1 - \widetilde{f}(s)}{s(1 - q\widetilde{f}(s))},$$

which is a formal solution to our problem in terms of the Laplace transforms. Note that it can be usually inverted only numerically.

3.8 Spatial Extreme Shock Model

In this section, we consider a two-dimensional model of spatial survival [10, 12]. It is a meaningful generalization of the univariate extreme shock model to the case of the spatial Poisson process of shocks. The random obstacles along the route of a moving object will play the role of these shocks. Although the initial setting is bivariate, the constructed failure rate is an univariate function and, therefore, our previous one-dimensional results can be used.

The setting of the problem is as follows: a sufficiently small normally or tangentially oriented interval is moving along a fixed route in the plane, crossing points of the spatial Poisson random process. Each crossing leads to a termination of the process (failure, accident) with a predetermined probability. As previously, the probability of passing the route without termination is of interest. An immediate application of the method to be considered is the *safety at sea* assessment. Our approach takes into account the fixed obstacles (e.g., shallows), which can lead to foundering and the moving obstacles (e.g., other ships), which can lead to collisions. The latter setting is not considered in this section and can be found in Finkelstein [12].

The field of fixed obstacles is considered to be random. In this application, there are two types of fixed obstacles: obstacles with known coordinates, marked in the corresponding navigational sea charts (and, therefore, not random), and obstacles with unknown coordinates, which following the subjective approach can be considered random. It turns out that, owing to the accuracy of navigation and motion control systems of a ship, weather influences, currents, etc., the obstacles with the known coordinates can also be modeled as random points in the plane. The 'geometric densities' of these obstacles, which can be obtained from the navigational charts, define the rates of the corresponding planar point processes to be used in the model [12].

The values of probabilities of accidents in "safety at sea" analysis are usually in the range 10^{-4} to 10^{-6}. Such estimates are often meaningless since there are not enough data to justify them. Therefore, simple relations for comparison of these probabilities can be very helpful in practice.

The developed approach can also be used for obtaining solutions that are optimal, for example, for finding a route with maximal probabilities of safe performance with or without specific restrictions (time on the route, fuel consumption, etc.). In what follows we consider the two-dimensional setting, but the generalization to $n = 3$ is straightforward and can be applied to assessing air traffic safety.

Denote by $\{N(B)\}$ an orderly point process in the plane, where $N(B)$ is a number of points in some domain $B \subset \Re^2$. We shall consider points of the process as prospective point influences (shocks) on our system (shallows for a ship, for instance). Similar to (2.12), the rate of this process $\lambda_f(\xi)$ can be formally defined as

$$\lambda_f(\xi) = \lim_{S(\delta(\xi)) \to 0} \frac{E[N(\delta(\xi))]}{S(\delta(\xi))}, \tag{3.41}$$

where $B = \delta(\xi)$ is the neighborhood of ξ with the area $S(\delta(\xi))$ and the diameter tending to zero. The subscript f stands for "fixed" obstacles.

Definition 3.3 The spatial nonhomogeneous Poisson process is defined similar to the one-dimensional case by the following relations [7]:

$$P\big(N(\delta(\xi)) = 1 | H_{\delta(\xi)}\big) = \lambda_f(\xi) S(\delta(\xi)) + o(S(\delta(\xi))),$$

$$P\big(N(\delta(\xi)) > 1 | H_{\delta(\xi)}\big) = o(S(\delta(\xi))),$$

where $H_{\delta(\xi)}$ denotes the configuration of all points outside $\delta(\xi)$.

It can be shown for an arbitrary B that $N(B)$ has a Poisson distribution with mean

$$\int_B \lambda_f(\xi) d\xi$$

and that the numbers of points in nonoverlapping domains are mutually independent random variables [7].

Our goal is to obtain a generalization of Eq. (3.18) to the bivariate case. The main feature of this generalization is a suitable parameterization allowing us to reduce the problem to the one-dimensional case [12]. Assume for simplicity that $\lambda_f(\xi)$ is a continuous function of ξ in an arbitrary closed circle in $\Re^2$. Let R_{ξ_1,ξ_2} be a fixed continuous curve connecting two distinct points in the plane, ξ_1 and ξ_2. We will call R_{ξ_1,ξ_2} a route. A point (a ship in our application) is moving in one direction along the route. Every time it 'crosses the point' of the process $\{N(B)\}$ (see later the corresponding regularization), an accident (failure) can happen with a given probability. We are interested in assessing the probability of moving along R_{ξ_1,ξ_2} without accidents. Let r be the distance from ξ_1 to the current point of the route (coordinate) and $\lambda_f(r)$ denote the corresponding rate. Thus, the one-dimensional parameterization is considered. For defining the corresponding

Poisson measure, the dimensions of objects under consideration should be taken into account.

Let $\left(\gamma_n^+(r), \gamma_n^-(r)\right)$ be a small interval of length $\gamma_n(r) = \gamma_n^+(r) + \gamma_n^-(r)$ in a normal direction to R_{ξ_1,ξ_2} at the point with the coordinate r, where the upper index denotes the corresponding direction ($\gamma_n^+(r)$ is on one side of R_{ξ_1,ξ_2}, whereas $\gamma_n^-(r)$ is on the other). Let $\overline{R} \equiv |R_{\xi_1\xi_2}|$ be the length of R_{ξ_1,ξ_2} and assume that the interval is small compared with the length of the route, i.e.,

$$\overline{R} > \; > \gamma_n(r), \forall r \in [0, \overline{R}].$$

The interval $\left(\gamma_n^+(r), \gamma_n^-(r)\right)$ is moving along R_{ξ_1,ξ_2}, crossing points of a random field. For "safety at sea" applications, it is reasonable to assume the symmetrical $\left(\gamma_n^+(r) = \gamma_n^-(r)\right)$ structure of the interval with length $\gamma_n(r) = 2\delta_s + 2\delta_o(r)$, where $2\delta_s$, $2\delta_o(r)$ are the diameters of the ship and of an obstacle, respectively. For simplicity, we assume that all obstacles have the same diameter. Thus, the ship's dimensions are already 'included' in the length of our equivalent interval. There can be other models as well, e.g., the diameter of an obstacle can be considered a random variable.

Taking Eq. (3.41) into account, the *equivalent rate* of occurrence of points, $\lambda_{e,f}(r)$ is defined as

$$\lambda_{ef}(r) = \lim_{\Delta r \to 0} \frac{E[N(B(r, \Delta r, \gamma_n(r)))]}{\Delta r}, \tag{3.42}$$

where $N(B(r, \Delta r, \gamma_n(r))$ is the random number of points crossed by the interval $\gamma_n(r)$ when moving from r to $r + \Delta r$. Thus, the specific domain in this case is defined as an area covered by the interval moving from r to $r + \Delta r$.

When $\Delta r \to 0$, $\gamma_n(r) \to 0$, and taking into account that $\lambda_f(\xi)$ is a continuous function [12],

$$\begin{aligned} E[N(B(r, \Delta r, \gamma_n(r)))] &= \int\limits_{B(r,\Delta r,\gamma_n(r))} \lambda_f(\xi) dS(\delta(\xi)) \\ &= \gamma_n(r)\lambda_f(r)\mathrm{d}r[1 + o(1)], \end{aligned}$$

which leads to the relationship for the equivalent rate of the corresponding one-dimensional nonhomogeneous Poisson process, i.e.,

$$\lambda_{ef}(r) = \gamma_n(r)\lambda_f(r)[1 + o(1)], \quad \Delta r \to 0, \; \gamma_n(r) \to 0.$$

As the radius of curvature of the route $R_c(r)$ is sufficiently large compared with $\gamma_n(r)$, i.e.,

$$\gamma_n(r) \ll R_c(r),$$

the domain covered by the interval $\left(\gamma_n^+(r), \gamma_n^-(r)\right)$ when it moves from r to $r + \Delta r$ along the route, is asymptotically ($\Delta r \to 0$) rectangular with area $\gamma_n(r)\Delta r$. Hence, the performed r -parameterization along the fixed route reduces the problem to the one-dimensional setting.

Assume now, as in the previous sections of this chapter, that the crossing of a point with a coordinate r leads to an accident (termination) with probability $p_f(r)$ and to the survival with the complementary probability $q_f(r) = 1 - p_f(r)$. Denote by R the random distance from the initial point of the route ξ_1 to a point of the route where an accident has occurred. Similar to (3.18), the probability of passing the route R_{ξ_1,ξ_2} without accidents can be derived in the following way:

$$P(R > \overline{R}) = \exp\left\{ -\int_0^{\overline{R}} \lambda_{af}(r)\mathrm{d}r \right\}, \tag{3.43}$$

where

$$\lambda_{af}(r) \equiv \theta_f(r)\lambda_{ef}(r) \tag{3.44}$$

is the corresponding failure (accident) rate. As previously, Eq. (3.43) and (3.44) constitute a simple and convenient tool for obtaining probabilities of safe (reliable) performance of our object. Thus, the univariate extreme shock model can be effectively applied to this initially two-dimensional setting.

3.9 Shock-Based Theory of Biological Aging

As a remarkable application to health sciences, we will show how the extreme shock model 'works' for obtaining the law of mortality of human populations. For this reason, we discuss and generalize the famous result by Strehler and Mildvan [29]. Our reasoning will mostly follow Finkelstein [15]. In this section, in accordance with the demographic and actuarial terminology, we will use the term "the force of mortality" (mortality rate) instead of the failure rate.

The Strehler–Mildvan [29] model suggests the justification of an exponential increase in the force of mortality $\mu(t)$, and describes some formal properties of the Gompertz mortality curve [17]:

$$\mu(t) = ae^{bt}. \tag{3.45}$$

The conventional generalization is the Gompertz–Makeham model, which adds a constant term c to the right-hand side of (3.45) in order to account for the 'background' mortality. In the current section, as in the original publication, we will assume that this term is negligible. Equation (3.45) usually provides a satisfactory fit to human mortality data for ages since maturity to the upper limit of around 90–100 years.

The goal of this section is to *discuss the underlying assumptions* of the Strehler–Mildvan (SM) *shock* model and the SM-correlation, which defines a negative correlation between parameters a and b. For several decades, the SM-correlation was believed to be a universal demographic law valid both for period and cohort mortality data [32].

The SM-model relies on the notion of vitality, i.e., an organism is characterized by its vitality function $V(t),\ V(0) \equiv V_0$, which decreases with age t. In the rest of this book, we will come back several times to the notion of vitality or its equivalents and will suggest a more mathematically advanced modeling of the vitality-related problems. Specifically, several strength–stress models will be considered when the failure (death) occurs if the magnitude of the stress (shock) exceeds the value of the strength (vitality).

According to Strehler and Mildvan [29], an organism is subject to stresses of internal or external nature that cause demands for energy. Those are shocks in our terminology. Let $(T_i,\ Y_i), i = 1, 2, \ldots$ be the sequence of pairs of i.i.d. random variables (therefore, the notation will be $(T,\ Y)$), characterizing the times at which stress events (demands for energy) occur, and the value of the demand for energy that is needed to recover from these stresses, respectively. Let $K(t)$ be the rate of the corresponding counting process describing arrival times of stress events. The following assumptions were made in the original paper:

Assumption 1 $\mathrm{Y_i}$ are exponentially distributed:

$$P(Y > y) = e^{-\frac{y}{D}}, \tag{3.46}$$

where, D is the mean value of this demand.

Assumption 2 An organism is characterized by its vitality function $V(t),\ V(0) \equiv V_0$ which decreases with age t. Yashin et al. [33], as in the original paper, called this function the maximum capacity of energy supply for an organism at age t. It can be also obviously interpreted as the stress resistance of an organism. Death occurs at age t when, for the first time, $Y > V(t)$. We discuss this assumption in conjunction with the last one.

Assumption 3 The rate $K(t) = K$ is a constant and the force of mortality is defined as [compare with Eq. (3.18)]

$$\mu(t) = KP(Y > V(t)) = Ke^{-\frac{V(t)}{D}}. \tag{3.47}$$

Equation (3.47) is called "a postulate" in Strehler and Mildvan [29]. However, it follows from the theory of point processes that (3.47) (see Chap. 2 and Sect. 3.4) is true *only* when the underlying point process $\{T_i\}_{i \geq 1}$ is the homogeneous Poisson process and, therefore, that the inter-arrival times of events (stresses) are exponentially distributed. This is a rather stringent condition, which was not pointed out in the original and subsequent papers discussing the SM-model. It should also be noted that, while (3.47), similar to (3.18), can be generalized to the case of the *nonhomogeneous Poisson process* with the age-dependent rate $K(t)$, the Poisson property of the underlying process is crucial for the product in the right-hand side of (3.47).

The following remark should be also made: as the force of mortality is a population characteristic, the vitality $V(t)$ should also be understood in this sense. However, it is obviously introduced by Assumption 1 as an individual (stochastic) characteristic. Therefore, we cannot simply substitute it with the corresponding expectation, as the exponential function is not linear:

$$E[e^{-\frac{V(t)}{D}}] \neq e^{-\frac{E[V(t)]}{D}}.$$

Thus, while there are a few important deficiencies in the original formulation of the model, it *formally* leads to the justified in practice properties of mortality rates.

Now we are ready to equate (3.45) and (3.47). As in the original paper, we will show using elementary derivations that $V(t)$ is *linearly declining* with age. It should be noted that this 'shape' is in consensus with the current understanding of the decline in the essential biological markers and the corresponding data, at least, for the human middle-age span [16]. Thus

$$\mu(t) = ae^{bt} = Ke^{-\frac{V(t)}{D}} \tag{3.48}$$

and taking logarithms of both sides ($V(0) \equiv V_0$):

$$V(t) = V_0(1 - (b/V_0)t) = V_0(1 - Bt), \tag{3.49}$$

where formally, $B = b/\ln(K/a) = Db/V_0$, and this quantity is usually called the individual rate of aging (in contrast with the population rate of aging b). Substituting (3.49) into (3.48):

$$\mu(t) = ae^{bt} = Ke^{-\frac{V_0(1-Bt)}{D}} = Ke^{-\frac{V_0}{D}}e^{\frac{V_0 Bt}{D}} \tag{3.50}$$

and thus

$$a = Ke^{-\frac{V_0}{D}}; \quad b = V_0 B/D. \tag{3.51}$$

Comparing two equations for the force of mortality, we see the dependence between a and b (negative correlation): the larger a results in the smaller b. From (3.51), this dependence can be written as

$$\ln a = \ln K - \frac{1}{B}b, \tag{3.52}$$

which is known in the literature as *SM-correlation*. This correlation has been observed empirically in various human populations. It follows from (3.52) that

$$\ln \mu(t) = \ln a + bt = \ln K + b(t - 1/B)), \tag{3.53}$$

meaning that the logarithms of mortality rates for different populations (e.g., with different a) intersect in one point with coordinates $(\ln K,\ 1/B)$. This has been experimentally observed and reported in the literature, although some criticism and violations of this rule were also discussed (see e.g., [32, 33]).

At first sight, it seems intriguing that the SM-correlation, which is derived using some general, partially unjustified assumptions, complies with the real mortality data. However, recently a certain departure from this pattern has been observed. A possible explanation is in consideration of the *vitality-independent* approach. It is based on the concept of lifesaving: i.e., that the environment not only supplies additional energy under stress, but due to the crucial advances in healthcare in recent decades, saves lives that previously would have been lost. The stochastic 'lifesaving model' (with a discussion of necessary assumptions) was developed in Finkelstein [11, 12]. It should be noted that Vaupel and Yashin [31] assumed that there can be a finite number of lifesavings, whereas we are dealing with a random number of these events.

Consider a lifetime that is characterized by the force of mortality $\mu(t)$ and the corresponding Cdf $F(t)$. Assume that a stress event affecting an organism, which occurs in accordance with this Cdf at age t_1 is fatal with probability $p(t_1)$ and is 'cured' with probability $1-p(t_1)$. The next stress occurs at age $t_2 > t_1$ in accordance with the Cdf $(F(t+t_1)-F(t))\overline{F}(t_1)$ and it is fatal with probability $p(t_2)$ and 'is cured' with probability $1-p(t_2)$, etc. It should be noted that the decreasing in age vitality of an organism can be still part of this model, if we assume that $1-p(t)$ is a decreasing function of age. In this case, $1-p(t)$ has a meaning of probability that the magnitude of a stress is smaller then the value of vitality at age t (probability of survival under a single shock). Therefore, in accordance with the lifesaving model [11], the initial nonhomogeneous Poisson process of stress events with rate $\mu(t)$ is terminated (*i.e.*, each event terminates the process with probability $p(t)$ and 'harmless' with probability $1-p(t)$) and the Cdf of time to termination is characterized by the force of mortality $p(t)\mu(t)$. Thus, we again arrive at our extreme shock model (3.18)!

In order to explain the departures from the Srtehler–Mildvan correlation that were observed in recent decades, assume now that probability $p(t)$ in the described lifesaving model is not age-dependent any more, *i.e.*, $p(t) \equiv p$. Obviously, the state of an organism (vitality) can 'affect' this probability, However, today it is mostly defined by the new 'technical' abilities of treating, e.g., medical conditions that could not be treated before or performing medical operations that were not possible before. Therefore, we can consider this probability as approximately constant. Our assumption also means that the proportion of conditions that can be now cured does not depend on age. Thus, the resulting force of mortality $p\mu(t)$ follows the proportional hazards (PH) model. In order to illustrate our further reasoning, consider the following example. Let Eq. (3.45) define the baseline force of mortality for a developed country at, e.g., chronological time $x_b = 1950$. Then it can be modified for time $x > x_b$ to

$$\mu_\tau(t) = p_\tau a e^{bt}, \tag{3.54}$$

where $\tau = x - x_b$ and p_τ is constant in age for the fixed τ. Thus, the environment, due to lifesaving and in accordance with the extreme shock model, 'decreases' only parameter a without affecting the slope of the logarithmic mortality rate b. This perfectly complies with the Gompertz shift model of Bongaarts and Feeney [4] and with other experimental studies. It also can explain the change in the rectangularization pattern (that is usually attributed to the Strehler–Mildvan correlation) to shifts in the corresponding survival curves (which can be explained by the PH model). The mortality data for developed countries in recent decades support these claims. It should be noted that the assumption of the underlying Gompertz law is essential for the described change in the pattern, which can be easily seen from Eq (3.54), as $p_\tau = e^{\ln p_\tau}$ ($\ln p_\tau < 0$) creates shifts in age for the baseline mortality rate. It is also worth mentioning that, although the method of constructing the resulting force of mortality in the SM model, which is captured by Eq (3.47), formally resembles our lifesaving approach, the difference lies in the fact that the corresponding probabilities are 'applied' to each stress event (with a constant rate) in the former case and to events occurring in accordance with the nonhomogeneous Poisson process with rate $\mu(t)$, in the latter case.

References

1. Bae S, Kuo W, Kvam P (2007) Degradation models and implied lifetime distributions. Reliab Eng Syst Saf 92:601–608
2. Barlow RE, Proschan F (1975). Statistical theory of reliability and life testing. Holt, Renerhart & Winston, New York
3. Beichelt FE, Fatti LP (2002) Stochastic processes and their applications. Taylor and Francis, London
4. Bongaarts J, Feeney G (2002) How long do we live? Popul Dev Rev 28:13–29
5. Cha JH, Finkelstein M (2009) On a terminating shock process with independent wear increments. J Appl Probab 46:353–362
6. Cha JH, Finkelstein M (2010) Burn-in by environmental shocks for two ordered subpopulations. Eur J Oper Res 206:111–117
7. Cox DR, Isham V (1980) Point processes. University Press, Cambridge
8. Cox DR, Miller HD (1965) Theory of stochastic processes. Methuen & Co, London
9. Dufresne F, Gerber H, Shiu E (1991) Risk theory with the gamma process. ASTIN Bull 21(2):177–192
10. Finkelstein M (1998) A point process stochastic model with application to safety at sea. Reliab Eng Syst Saf 60:227–234
11. Finkelstein M (2005) Lifesaving explains mortality decline with time. Math Biosci 196:187–197
12. Finkelstein M (2008) Failure rate modelling for reliability and risk. Springer, London
13. Finkelstein M, Marais F (2010) On terminating Poisson processes in some shock models. Reliab Eng Syst Saf 95:874–879
14. Finkelstein M, Zarudnij VI (2002) Laplace transform methods and fast repair approximations for multiple availability and its generalizations. IEEE Trans Reliab 51:168–177
15. Finkelstein M (2012) Discussing the Strehler-Mildvan model of mortality. Demographic Res. 26:191–206
16. Golubev A (2009) How could the Gompertz-Makeham law evolve. J Theor Biol 238:1–17

17. Gompertz B (1825) On the nature of the function expressive of the law of human mortality and on a new mode of determining the value of life contingencies. Philos Trans R Soc 115:513–585
18. Gut A (1990) Cumulated shock models. Adv Appl Probab 22:504–507
19. Gut A, Husler J (2005) Realistic variation of shock models. Stat Probab Lett 74:187–204
20. Kahle W, Wendt H (2004) On accumulative damage process and resulting first passage times. Appl Stoch Models Bus Ind 20:17–27
21. Kalashnikov V (1997) Geometric sums: bounds for rare events with applications. Kluwer Academic Publishers, Dordrecht
22. Lam Y, Zhang YL (2004) A shock model for the maintenance problem of a repairable system. Comput Oper Res 31:1807–1820
23. Lund R, McCormic W, Xiao U (2004) Limiting properties of Poisson shot noise processes. J Appl Probab 41:911–918
24. Papoulis A (1971) High density shot noise and Gaussianity. J Appl Probab 8:118–127
25. Rice J (1977) On generalized shot noise. Adv Appl Probab 9:553–565
26. Ross SM (1996) Stochastic processes, 2nd edn. Wiley, New York
27. Singpurwalla N (1995) Survival in dynamic environment. Stat Sci 10:86–103
28. Tang YY, Lam Y (2006) A delta-shock maintenance model for a deteriorating system. Eur J Oper Res 168:541–556
29. Strehler L, Mildvan AS (1960) General theory of mortality and aging. Science 132:14–21
30. Van Nortwijk JM (2009) A survey of the application of gamma processes in maintenance. Reliab Eng and Syst Saf 94:2–21
31. Vaupel JW, Yashin AI (1987) Repeated resuscitation: how life saving alters life tables. Demography 4:123–135
32. Yashin AI, Begun AS, Boiko SI, Ukraintseva SV, Oeppen J (2001) The new trends in survival improvement require a revision of traditional gerontological concepts. Exp Gerontol 37:157–167
33. Yashin AI, Ukraintseva SV, Boiko SI, Arbeev KG (2002) Individual aging and mortality rate: how are they related. Soc Biol 49:206–217