7. Occurrence Frequency Distributions

> *Probability is expectation founded upon partial knowledge. A perfect acquaintance with all the circumstances affecting the occurrence of an event would change expectation into certainty, and leave neither room nor demand for a theory of probabilities.*
>
> George Boole (1815–1864)

> *Nevertheless, as is a frequency occurrence in science, a general hypothesis was constructed from a specific instances of a phenomenon.*
>
> Sidney Altman (born 1939)

It is customary in the statistics of nonlinear processes to histogram the logarithmic number of events versus a logarithmic size scale, which is called a *log* $N -$ *log* S *diagram*, *size distribution*, *occurrence frequency distribution*, or simply *frequency distribution*. In such log-log representations, the difference between (i) a Poissonian random process, which can be characterized by an exponential distribution function that drops off sharply above an e-folding size scale, and (ii) nonlinear processes governed by self-organized criticality, which ideally produce a scale-free powerlaw distribution function, appears most striking. Frequency distributions thus have become the arbiters of SOC versus non-SOC processes, starting from the famous magnitude diagram of earthquakes discovered by Beno Gutenberg and Charles Francis Richter in 1954 (i.e., the Gutenberg–Richer law). Frequency distributions of SOC phenomena obtained from astrophysical data were first identified in solar flare data by Ed Lu and Russell Hamilton (1991), based on log-log histograms published earlier without an interpretation in terms of SOC (e.g., Dennis 1985). Frequency distributions can be plotted for any conceivable parameter, preferably a (model-independent) observable, which does not require an arbitrary choice of a physical model. For earthquake statistics, the most commonly used parameter is the magnitude, measured by well-calibrated seismometers. In astrophysical data, where the observable is typically a time series of flux intensity in some given wavelength range, obvious parameters used for frequency distributions are the peak flux P, the total (time-integrated) flux or fluence E, and the total duration T of an event. While such observables can be unambiguously measured from well-calibrated detectors, the ultimate desire is to obtain frequency distributions of physical parameters, such as the thermal energy E_{th}, the nonthermal energy E_{nth}, the magnetic energy E_B, the kinetic energy E_{kin}, or the potential energy E_{pot}, which of course are all model-dependent, to be discussed in Chapter 9. In the following section we review the occurrence frequency distributions of (preferably) observables measured from astrophysical events that are (hypothetically) associated with SOC processes and we will compare the observations with analytical SOC models (Chapter 3) in order to evaluate their consistency with SOC theory.

7.1 Basics of Frequency Distribution Functions

The data input for an occurrence frequency distribution is usually a list or a catalog of events, characterized by some size parameter x_i for $i = 1, ..., n$ events, regardless whether the list was generated by visual inspection or by an automated computer algorithm (Chapter 6). How do we construct a log-log histogram from an event catalog? There are essentially two ways, either a logarithmically binned histogram if large statistics is available, or a rank-order plot if the size of the statistical sample is rather small.

7.1.1 Differential Frequency Distributions

If we have large statistics (at least $n \gtrsim 10^2, ..., 10^3$), we can first establish a logarithmic binning axis bound between the minimum and maximum value, $x_{min} \le x_i \le x_{max}$, with n_j bins,

$$x_j^{bin} = x_{min} \left(\frac{x_{max}}{x_{min}} \right)^{(j-1)/(n_j-1)} , \qquad j = 1, ..., n_j \tag{7.1.1}$$

which is equidistant on a logarithmic scale, but has variable intervals on a linear scale,

$$\Delta x_j^{bin} = x_{j+1}^{bin} - x_j^{bin} = x_j^{bin} \left[\left(\frac{x_{max}}{x_{min}} \right)^{1/(n_j-1)} - 1 \right] . \tag{7.1.2}$$

In a next step we can count the number of events N_j^{bin} that fall in each bin with interval $x_j^{bin} \le x_i \le x_{j+1}^{bin}$. We have to be aware that this number N_j^{bin} depends on the particular bin size Δx_j^{bin} we have chosen. In order to obtain the functional form of the frequency distribution, which should be independent of the binning, we have to divide the number of events by the bin size,

$$N_j = N(x_j) = \frac{N_j^{bin}}{\Delta x_j^{bin}} , \tag{7.1.3}$$

and can plot the frequency distribution with N_j on the y-axis versus the size x_j on the x-axis. This representation normalizes the distribution to the total number of events n, which we should obtain by integration over the x-axis, or summing over all (non-equidistant) bins on the x-axis,

$$\int_0^\infty N(x)\,dx = \sum_{j=0}^{n_j} N_j\,dx_j = \sum_{j=0}^{n_j} N(x_j)\Delta x_j^{bin} = \sum_{j=0}^{n_j} N_j^{bin} = n\,. \tag{7.1.4}$$

Instead of expressing the number of occurrences by the actually observed numbers N_j, it is also customary to use a *probability distribution function* $P(x)$, which is simply the number of events in each bin normalized by the total number of events n,

$$P(x) = \frac{N(x)}{n} \tag{7.1.5}$$

which has the total integral normalized to unity,

$$\int_{x_{min}}^{x_{max}} P(x)\,dx = \int_{x_{min}}^{x_{max}} \frac{N(x)}{n}\,dx = 1\,. \tag{7.1.6}$$

Both representations, $N(x)$ or $P(x)$, are called a *differential frequency distribution*, because they express the number of events in a “differential” bin dx.

7.1.2 Cumulative Frequency Distributions

An integrated *differential frequency distribution* $N(x)dx$ is called a *cumulative frequency distribution* $N^{cum}(>x)$, which expresses in each bin the sum of all events that are larger than the size parameter of the bin x,

$$N^{cum}(>x) = \int_x^{x_{max}} N(x)dx\,, \tag{7.1.7}$$

which we denote by $N^{cum}(>x)$, in contrast to the differential distribution $N(x)$. The cumulative frequency distribution contains more statistics in the rarer bins at larger sizes, and thus appears smoother at the upper end than differential distribution functions. However, the values in each bin are statistically not independent, but always contain information from all other bins on the right-hand side. The particular functional shape at the upper cutoff can dominate the entire distribution function.

If the differential frequency distribution is a powerlaw function with slope α, the cumulative frequency distribution is expected to have a flatter powerlaw slope by one,

$$\begin{array}{ll} N(x) & \propto x^{-\alpha} \\ N^{cum}(>x) & \propto x^{-\beta} \lesssim x^{-(\alpha-1)} \end{array}\,. \tag{7.1.8}$$

For instance, both the differential and cumulative frequency distributions for earthquakes are shown in the same plot (Fig. 1.7), with slopes of $\alpha = 2$ and $\beta = 1$. However, the powerlaw relationship $\beta = \alpha - 1$ is only true when the differential frequency distribution extends to infinite, which is never the case. In reality, there is always a largest event at x_{max},

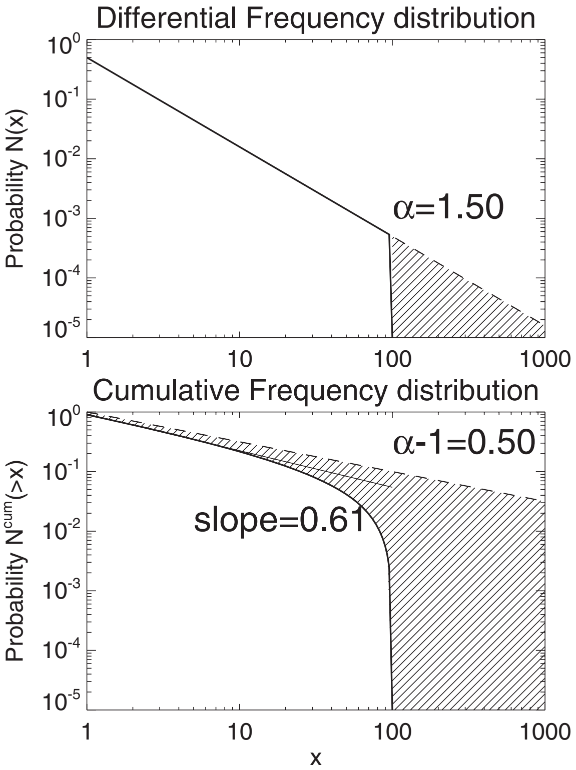

Fig. 7.1 A differential frequency distribution function $N(x)$ is shown with a powerlaw slope of $\alpha = 1.5$ and cutoff at $x_{max} = 100$ (top). The corresponding cumulative frequency distribution function $N^{cum}(>x)$ is fitted in the lower half (logarithmic) range $x = [1,...,10]$, which gives a powerlaw slope of $\beta = 0.61$, which is steeper than expected for an ideal powerlaw distribution without upper cutoff x_{max}, i.e., $\beta = \alpha - 1 = 0.5$.

which causes a sharp cutoff in the differential frequency distribution $N(x)$, but a gradual steepening in the cumulative frequency distribution, because of the missing contributions from $x > x_{max}$. This detail is quite important, because it leads to a significant over-estimate of the powerlaw slope when the rule $\alpha = \beta + 1$ is applied. We demonstrate this in the following. We define a powerlaw distribution function with a sharp cutoff at x_{max},

$$N(x) = (\alpha - 1) x^{-\alpha} \ , \quad x \le x_{max} \ . \tag{7.1.9}$$

The cumulative frequency distribution function can then be calculated by integrating over the range from x to x_{max},

$$N^{cum}(>x) = n \frac{\int_x^{x_{max}} N(x')\, dx'}{\int_{x_{min}}^{x_{max}} N(x')\, dx'} = n \frac{\int_x^{x_{max}} x'^{-\alpha}\, dx'}{\int_{x_{min}}^{x_{max}} x'^{-\alpha}\, dx'} = n \frac{(x^{1-\alpha} - x_{max}^{1-\alpha})}{(x_{min}^{1-\alpha} - x_{max}^{1-\alpha})} \ . \tag{7.1.10}$$

We see that the second term in the integral, $x_{max}^{1-\alpha}$ steepens the slope and lets the cumulative distribution drop to zero when x approaches x_{max}. We plot the two distribution functions in Fig. 7.1 for $\alpha = 1.5$. The powerlaw slope of the cumulative distribution function is expected to be $\beta = \alpha - 1 = 1.5 - 1 = 0.5$ at the lower end $x \ll x_{max}$, but becomes systematically steeper near the upper cutoff. If we were to fit a powerlaw over the powerlaw-like range of x, say in the range $x_{min} \le x \le x_{max}/10$, we would measure a slope of $\beta = 0.61$ (Fig. 7.1). This steepening effect on the slope due to the presence of an upper cutoff does not occur in the differential distribution, so it is important to take this effect into consideration when dealing with cumulative frequency distribution functions, e.g., see Fig. 1.7 for earthquakes or Fig. 1.15 for stellar flares.

How can this upper cutoff effect be taken properly into account? The best way is to fit the exact analytical function of the cumulative frequency distribution function, which is $N^{cum}(>x) \propto (x^{1-\alpha} - x_{max}^{1-\alpha})$ (Eq. 7.1.10), rather than the powerlaw approximation $N^{cum}(>x) \propto x^{1-\alpha}$. Alternatively, if the original data are not available, but only a powerlaw fit to the cumulative distribution is known (e.g., from literature),

$$N^{cum}(>x) = n \left(\frac{x}{x_{min}} \right)^{-\beta} , \tag{7.1.11}$$

we can calculate the relationship between the cumulative powerlaw slope β and the differential powerlaw slope α. A practical way is to assume that the cumulative powerlaw slope β gives a good fit of the parameter x in the lower half (logarithmic) range $[x_{min}, x_{max}]$ (see Fig. 7.1), which we define with the fractions $[q_1, q_2]$ with $q_1 = x_{min}/x_{max}$ and $q_2 = q_1^{1/2}$. For instance, for $x_{min} = 1$ and $x_{max} = 100$ (Fig. 7.1), the lower logarithmic half has the fractions $q_1 = 0.01$ and $q_2 = 0.1$. From the cumulative powerlaw fit we have the following occurrence ratio between these two points (using $q_1 = q_2^2$),

$$\frac{N^{cum}(>x_2)}{N^{cum}(>x_1)} = \left(\frac{x_2}{x_1} \right)^{-\beta} = \left(\frac{q_2}{q_1} \right)^{-\beta} = q_2^{\beta} . \tag{7.1.12}$$

On the other side, from Eq. (7.1.10) we have, using $q_1 = q_2^2$ and applying $(x^2 - 1) = (x-1)(x+1)$ for $x = q_2^{1-\alpha}$,

$$\frac{N^{cum}(>x_2)}{N^{cum}(>x_1)} = \left(\frac{q_2^{1-\alpha} - 1}{q_1^{1-\alpha} - 1} \right) = \left(\frac{q_2^{1-\alpha} - 1}{q_2^{2(1-\alpha)} - 1} \right) = \frac{1}{q_2^{(1-\alpha)} + 1} . \tag{7.1.13}$$

Setting these two expressions (Eqs. 7.1.12 and 7.1.13) equal, we obtain the following relationship for α as a function of β,

$$\alpha = 1 - \frac{\log[q_2^{-\beta} - 1]}{\log(q_2)} . \tag{7.1.14}$$

or vice versa, the relationship for β as a function of α,

$$\beta = -\frac{\log [q_2^{(1-\alpha)} + 1]}{\log(q_2)} \ . \qquad (7.1.15)$$

We see that both expressions yield the approximation $\alpha \approx 1+\beta$, if $q_2^{-\beta} \gg 1$ or $q_2^{1-\alpha} \gg 1$, which comes down to the condition of large logarithmic ranges, i.e. $x_{max} \gg x_{min}$. For the case shown in Fig. 7.1 with $q_2 = 0.1$ and $\alpha = 1.5$, we obtain $\beta = \log(1+0.1^{-0.5}) = 0.62$, which is significantly different from the approximation $\beta \approx \alpha - 1 = 0.5$. The difference is even larger for smaller logarithmic ranges, say for one decade ($x_{min}/x_{max} = 0.1$), as it is the case for small samples, such as statistics of stellar flares (Fig. 1.15).

7.1.3 Rank-Order Plots

If the statistical sample is rather small, in the sense that no reasonable binning of a histogram can be done, either because we do not have multiple events per bin or because the number of bins is too small to represent a distribution function, we can create a rank-order plot. A rank-order plot is essentially an optimum adjustment to minimum statistics that gives every single event a single bin. From an event list of a parameter $x_i, i = 1,...,n$, which is generally not sorted, we have first to generate a rank-ordered list by ordering the events according to increasing size,

$$x_1 \le x_2 \le ... \le x_j \le ... \le x_n \ , \quad j = 1,...,n \ . \qquad (7.1.16)$$

The bins are generally not equidistant, neither on a linear nor logarithmic scale, defined by the difference between subsequent values of the ordered x_j,

$$\Delta x_j^{bin} = x_{j+1}^{bin} - x_j^{bin} \ . \qquad (7.1.17)$$

In a rank-ordered sequence of n events, the probability for the largest value is $1/n$, for events that are larger than the second-largest event it is $2/n$, and so forth, while events larger than the smallest event occur in this event list with a probability of unity. Thus, the cumulative frequency distribution is simply the reversed rank order,

$$N^{cum}(>x_j) = (n+1-j) \ , \qquad j = 1,...,n \ , \qquad (7.1.18)$$

and the distribution varies from $N^{cum}(>x_1) = n$ for $j = 1$ to $N^{cum}(>x_n) = 1$ for $j = n$. We can plot a cumulative frequency distribution with $N^{cum}(>x_j)$ on the y-axis versus the size x_j on the x-axis. The distribution is normalized to the number of events n,

$$\int_{x_1}^{x_n} N(x)\, dx = N^{cum}(>x_1) = n \ . \qquad (7.1.19)$$

The differential frequency distribution function $N(x)$ could in principle be computed from the derivative of the cumulative distribution, but there is usually considerable noise between subsequent events in a rank order, so that smoothing is recommended before differentiation.

We show two examples of rank-ordered plots in Fig. 7.2. The first example is based on a differential frequency distribution of time scales that correspond to an exponential function with time scale $\tau = 1$,

$$N(t) = \frac{1}{\tau} \exp\left(-\frac{t}{\tau}\right) , \tag{7.1.20}$$

Using a random generator we are producing $n = 100$ values of time scales $t_i, i = 1, \ldots, n$ that correspond to this differential distribution according to the method described in the following Section 7.1.4, which we plot in a rank-ordered diagram as shown in Fig. 7.2 (left; diamonds). In order to prove that this rank-order plot corresponds to the cumulative distribution, we calculate the distribution analytically by integrating Eq. (7.1.20),

$$N^{cum}(>t) = \int_t^{\infty} N(t')\, dt' = \int_t^{\infty} \frac{1}{\tau} \exp\left(-\frac{t'}{\tau}\right) dt' = \exp\left(-\frac{t}{\tau}\right) . \tag{7.1.21}$$

which agrees (Fig. 7.2 left, thick solid curve) with the rank-ordered values.

The second example is based on a differential frequency distribution of energies that have a powerlaw function with slope of $\alpha = 1.5$,

$$N(E) = (\alpha - 1) E^{-\alpha} . \tag{7.1.22}$$

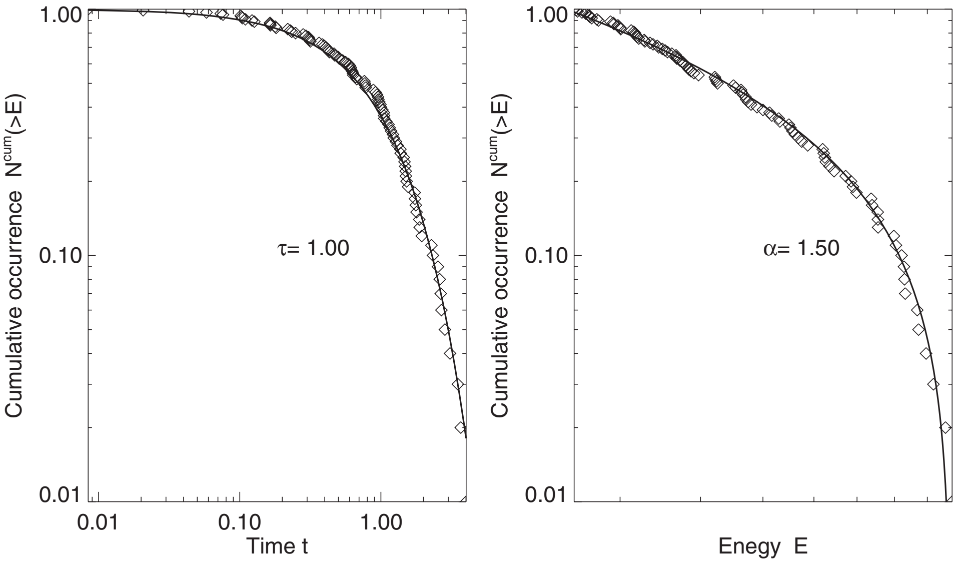

Fig. 7.2 Numerically generated rank-order plots (diamonds) and theoretical cumulative frequency distribution functions (thick solid curves) for an exponential function with $\tau = 1$ (left) and a powerlaw function with $\alpha = 1.5$ (right).

Again, using a random generator we are producing $n = 100$ values of energy $E_i, i = 1, ..., n$ that correspond to this differential distribution according to the method described in the following Section 7.1.4, which we plot in a rank-order diagram as shown in Fig. 7.2 (right; diamonds). In order to prove that this rank-order plot corresponds to the cumulative distribution, we calculate the distribution analytically by integrating Eq. (7.1.22),

$$N^{cum}(>E) = \int_E^{E_{max}} (\alpha - 1)\varepsilon^{-\alpha}\, d\varepsilon = E^{(1-\alpha)} - E_{max}^{1-\alpha} . \tag{7.1.23}$$

which agrees (Fig. 7.2 right, thick solid curve) with the rank-ordered values.

An observational example of cumulative frequency distributions based on a rank-order plot is shown in Fig. 1.15 for stellar flares, where the statistics literally does not include more than about a dozen events per star (Audard et al. 2000).

Sometimes it is also convenient to plot the size versus the rank, such as the ranking of cities by population size shown in Fig. 1.4. This is essentially the rank-order plot defined in Eq. (7.1.18), but with exchanged x- and y-axis. The rank order on the x-axis is the independent variable N_j, while the y-axis is the dependent variable $x_j = x(N_j)$. Since the axes are exchanged, a powerlaw function would have approximately the reciprocal value for the slope,

$$\begin{aligned} N(x) &\propto x^{-\alpha} \\ x(N) &\approx N^{-1/\alpha} \end{aligned} \tag{7.1.24}$$

This type of rank-order plot with size versus rank was originally used by Zipf (1949) for statistics of word usage (Section 1.3), and thus is also called *Zipf plot*.

7.1.4 Numerical Generation of Frequency Distributions

For numerical simulations of frequency distributions, for instance Monte-Carlo simulations of SOC models, we need to create randomly distributed values x_i that have a particular prescribed function of their frequency distribution, such as an exponential function for waiting times, or a powerlaw function for energies. Let us prescribe the form of the frequency distribution with a probability function $p(x)$ in the interval $[x, x+dx]$, which has the normalization,

$$\int_0^\infty p(x)\, dx = 1 . \tag{7.1.25}$$

The total probability $\rho(x)$ to have a value in the range of $[0, x]$ is then the integral,

$$\rho(x) = \int_0^x p(x')\, dx' . \tag{7.1.26}$$

If the analytical function $\rho(x)$ can be inverted, say with the analytical inverse function ρ^{-1}, so that

$$x = \rho^{-1}(\rho) = \rho^{-1}(\rho[x]) , \tag{7.1.27}$$

we have a transformation that allows us to generate values x_i from a distribution of probability values ρ_i. There are many numerical random generator algorithms available that produce a random number ρ_i in a unity range of $[0,1]$, which can then be used to generate values x_i with the mapping transform $x_i = \rho^{-1}(\rho_i)$. The frequency distribution of these values x_i will then fulfill the prescribed function $p(x)$.

As an example we demonstrate the numerical generation of a sample of time scales t that has a frequency distribution function following an exponential function,

$$p(t) = \frac{1}{\tau} \exp\left(-\frac{t}{\tau}\right) , \tag{7.1.28}$$

which fulfills the normalization $\int_0^1 p(t)\, dt = 1$. The total probability $\rho(t)$ to have a value in the range $[0,t]$ is then the integral function of $p(t)$,

$$\rho(t) = \int_0^t p(t')\, dt' = \int_0^t \frac{1}{\tau} \exp\left(-\frac{t'}{\tau}\right) dt' = \left[1 - \exp\left(-\frac{t}{\tau}\right)\right] . \tag{7.1.29}$$

The inverse function $t(\rho)$ of $\rho(t)$ is

$$t(\rho) = -\tau \ln(1-\rho) . \tag{7.1.30}$$

In Fig. 7.3 (left) we use a random generator that produces 10,000 values ρ_i, uniformly distributed in the range of $[0,1]$, and use the transform Eq. (7.1.30) to generate values

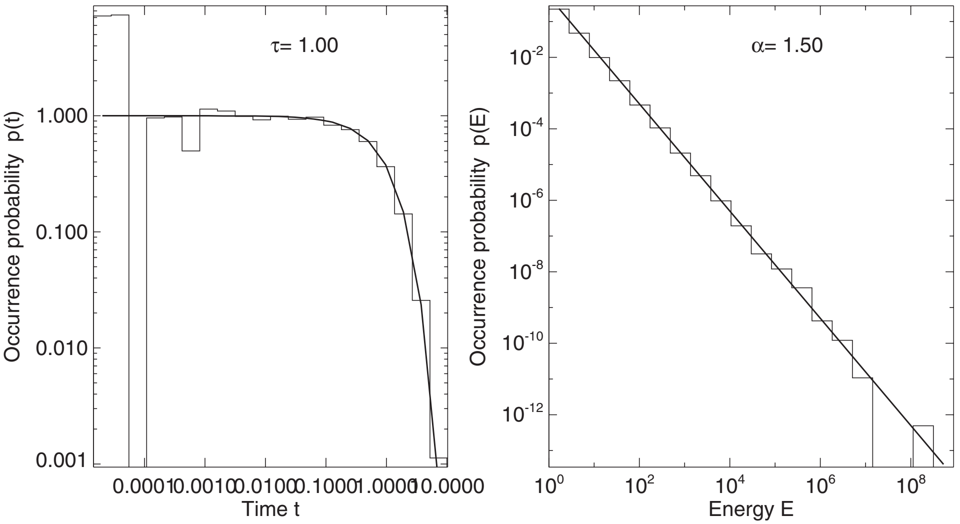

Fig. 7.3 *Left:* An exponential frequency distribution is numerically generated, based on $n = 10{,}000$ uniformly distributed values ρ_i in the range $[0,...,1]$ and times $t_i = -\tau \ln(1-\rho_i)$ (Eq. 7.1.30) (histogram), leading to the occurrence probability function $p(t)$ as defined in Eq. (7.1.28) (thick curve). *Right:* A powerlaw frequency distribution is numerically generated, based on $n = 10{,}000$ values ρ_i uniformly distributed in the range $[0,...,1]$, with energies $E_i = (1+\rho_i)^{1/(1-\alpha)}$ with $\alpha = 1.5$ (Eq. 7.1.33) (histogram), leading to the occurrence probability function $p(E)$ as defined in Eq. (7.1.31) (thick curve).

$t_i = -\tau \ln(1-\rho_i)$ with $\tau = 1$ and sample the frequency distribution of the 10,000 values t_i, which follows the prescribed exponential function $p(t)$ defined in Eq. (7.1.28).

As a second example we prescribe a powerlaw function $p(E)$ for the frequency distribution,

$$p(E) = (\alpha-1)E^{-\alpha} \ , \tag{7.1.31}$$

which fulfills the normalization $\int_1^\infty p(E)\, dE = 1$. The total probability $\rho(E)$ in the range $[0,E]$ is then the integral function of $p(E)$ (Eq. 7.1.31),

$$\rho(E) = \int_0^E p(\varepsilon)\, d\varepsilon = \int_0^E (\alpha-1)\varepsilon^{-\alpha}\, d\varepsilon = \left[1 - E^{(1-\alpha)}\right] \ . \tag{7.1.32}$$

The inverse function $E(\rho)$ of $\rho(E)$ (Eq. 7.1.32) is

$$E(\rho) = [1-\rho]^{1/(1-\alpha)} \ . \tag{7.1.33}$$

In Fig. 7.3 (right) we use a random generator that produces 10,000 values ρ_i uniformly distributed in the range of $[0,1]$, choose a powerlaw index of $\alpha = 1.5$, and use the transform Eq. (7.1.33) to generate values $E_i = [1-\rho_i]^{-2}$ and sample the frequency distribution of the 10,000 values E_i, which follows the prescribed powerlaw function $p(E) = 0.5E^{-1.5}$ as defined in Eq. (7.1.31).

7.1.5 Integrals of Powerlaw Distributions

For normalization purposes or when the total number n of events needs to be evaluated from a powerlaw distribution $N(x) = (\alpha-1)x^{-\alpha}$, we have to integrate over the valid range bound by $x_{min} < x < x_{max}$,

$$n = \int_{x_{min}}^{x_{max}} N(x)\, dx = \int_{x_{min}}^{x_{max}} (\alpha-1)x^{-\alpha}\, dx = x_{min}^{1-\alpha} - x_{max}^{1-\alpha} \ , \tag{7.1.34}$$

which is defined for $\alpha \neq 1$. Generally both boundaries contribute significantly to the total number, unless the powerlaw distribution extends over a very large range, say more than three orders of magnitude. For such large ranges, the following approximations can be used,

$$n = \begin{cases} \approx x_{min}^{1-\alpha} \text{ for } (x_{max} \gg x_{min}) \text{ and } (\alpha > 1) \\ \approx x_{max}^{1-\alpha} \text{ for } (x_{max} \gg x_{min}) \text{ and } (\alpha < 1) \end{cases} \ . \tag{7.1.35}$$

The total integral (or first moment) of a powerlaw distribution function, for instance the total energy of an occurrence frequency distribution of energies, can be obtained by convolving the variable x with the powerlaw distribution $N(x) = (\alpha-1)x^{-\alpha}$ over the valid range $x_{min} \le x \le x_{max}$,

$$x_{tot} = \int_{x_{min}}^{x_{max}} x\, N(x)\, dx = \int_{x_{min}}^{x_{max}} (\alpha-1)x^{1-\alpha}\, dx = \left(\frac{\alpha-1}{2-\alpha}\right) \left[x_{max}^{2-\alpha} - x_{min}^{2-\alpha}\right] \ , \tag{7.1.36}$$

which is only defined for $\alpha \neq 2$. Again, generally both boundaries contribute significantly to the total number, unless the powerlaw distribution extends over a very large range, in which case the following approximations can be used,

$$x_{tot} \approx \left(\frac{\alpha-1}{2-\alpha}\right) \begin{cases} x_{min}^{2-\alpha} & \text{for } (x_{max} \gg x_{min}) \text{ and } (\alpha > 2) \\ x_{max}^{2-\alpha} & \text{for } (x_{max} \gg x_{min}) \text{ and } (\alpha < 2) \end{cases} . \tag{7.1.37}$$

The critical value is $\alpha = 2$, which decides whether the integral diverges at the lower bound (if $\alpha > 2$) or upper bound (if $\alpha < 2$). A far-reaching application of this integral is the total energy contained in the distribution of solar or stellar flares, which is also responsible for heating of the solar or stellar corona, and could be dominated by nanoflares if $\alpha > 2$ applies over a large energy range, as pointed out by Hudson (1991).

7.1.6 Powerlaw Scaling Laws and Correlations

We consider the case where two parameters x and y are correlated with each other by a powerlaw function with the power coefficient β,

$$y \propto x^{\beta} , \tag{7.1.38}$$

where x and y could be the peak energy P, the total energy E, duration T, or any other SOC parameter. Since every SOC parameter has a powerlaw-like distribution function in our SOC standard model (Section 3.1),

$$\begin{aligned} N(x) &\propto x^{-\alpha_x} \\ N(y) &\propto y^{-\alpha_y} \end{aligned} , \tag{7.1.39}$$

it is useful to calculate the relationship between the power indices α_x, α_y, and β. The general way to substitute a variable $y(x)$ in a frequency distribution $N(x)$ is,

$$N(y)\, dy = N[x(y)] \left| \frac{dx(y)}{dy} \right| dy , \tag{7.1.40}$$

which yields, for the function $y(x) \propto x^{\beta}$ defined in Eq. (7.1.38), using the inverse function $x(y) \propto y^{1/\beta}$ and derivative $dx/dy \propto y^{(1/\beta-1)}$,

$$N(y)\, dy = y^{-\alpha_x/\beta+1/\beta-1}\, dy = y^{-\alpha_y}\, dy , \tag{7.1.41}$$

leading to the following relationship between the power indices,

$$\beta = \frac{(\alpha_x - 1)}{(\alpha_y - 1)} . \tag{7.1.42}$$

This is a useful relationship to compute (or verify) the power index β of two correlated parameters from their frequency distributions. For instance in Fig. 6.1, we have for $x = P$ and $y = E$ the frequency distributions with a powerlaw slope of $\alpha_x \approx 2.0$ and $\alpha_y \approx 1.5$, which predicts $\beta = (\alpha_x - 1)/(\alpha_y - 1) = (2.0 - 1)/(1.5 - 1) = 1/0.5 = 2$, which indeed corresponds to the power index $\beta \approx 2.0$ in the scatterplot of $E \propto P^2$ (Fig. 6.1, third row middle).

7.1.7 Accuracy of Powerlaw Fits

There are a number of effects that determine the accuracy of derived powerlaw indices in frequency distributions, $N(x) \propto x^{-\alpha}$, and underlying correlations, $y(x) \propto x^{\beta}$, such as formal errors of linear regression fits, the choice of dependent and independent variables, the statistical uncertainty of the number of events, a small logarithmic range, statistical weighting, histogram binning, cutoffs, truncations, deviations from powerlaws, sensitivity limits, etc. In addition, the sampling of events is also affected by data pre-processing, such as dead-time corrections, spike removals, or background subtractions. When error bars are given in literature, they usually refer to formal errors of a least-squares fit, but do not include systematic errors that result from numerous biases of the data set or model assumptions. The best way to obtain a realistic error is often to use a Monte-Carlo simulation of the data and measurement procedure based on a realistic model of the data. In order to give a typical assessment of various errors in our study of occurrence frequency distributions and parameter correlations we conduct some Monte-Carlo simulations of our standard SOC model (Section 3.1).

Following the numerical simulation procedure outlined in Section 7.1.4 we generate two sets of $n = 100$ variables, one for peak energies $P_i, i = 1, ..., n$ and one for total energies $E_i, i = 1, ..., n$, which both have powerlaw distribution functions, with powerlaw slopes of $\alpha_P = 2$ and $\alpha_E = 1.5$,

$$\begin{aligned} N(P) &= (\alpha_P - 1)P^{-\alpha_P} \\ N(E) &= (\alpha_E - 1)P^{-\alpha_E} \end{aligned} \quad . \tag{7.1.43}$$

They can be generated using two sets of random numbers ρ_i and ρ_j uniformly distributed in the range $[0, ..., 1]$, produced with a numerical random generator and the transform (Eq. 7.1.33),

$$\begin{aligned} P_i &= [1 - \rho_i]^{1/(1-\alpha_P)} \\ E_j &= [1 - \rho_j]^{1/(1-\alpha_E)} \end{aligned} \tag{7.1.44}$$

In order to ensure a parameter correlation we sort each set in increasing number, but add some random noise in both parameters. We mimic also an instrumental sensitivity limit by applying a flux threshold of $P \geq 0.5$, which causes a truncation error in P.

We plot the two sets of variables in form of a scatterplot E_j versus P_i in Fig. 7.4 (top left), which show a linear regression fit of $E \propto P^{1.78}$ instead of the theoretically expected relationship $E \propto P^2$. A linear regression fit with inverted axis yields $P \propto E^{0.52}$, which corresponds to $E \propto P^{(1/0.52)} = P^{1.93}$, so part of the difference results from the choice of the independent variable. There are other linear regression fits that treat both variables equally, such as the bisector method or the minimization of the orthogonal distance to the linear regression fit, which eliminate this bias. However, there is still an additional bias introduced by the flux threshold, which affects a truncation of data for P but not for E. This could be corrected by using a truncation limit that is orthogonal to the linear regression fit.

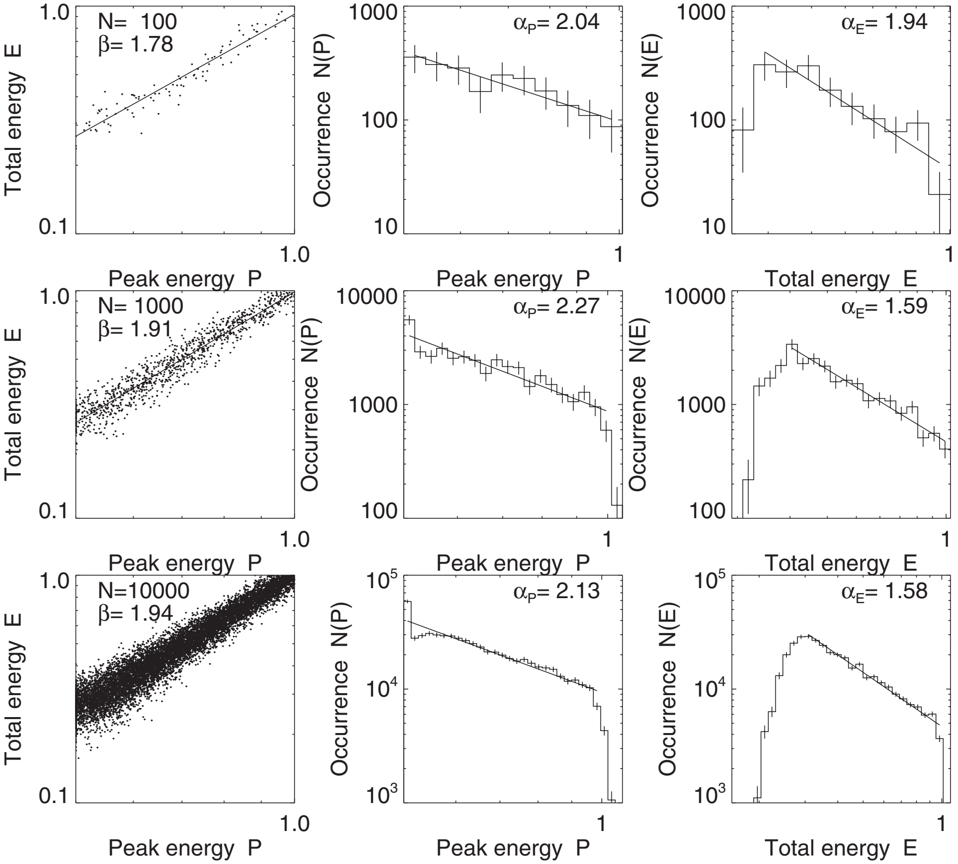

Fig. 7.4 Three Monte-Carlo simulations of peak energies P (middle column) and total energies E (right column) from random samples of prescribed powerlaw distributions $N(P) \propto P^{2.0}$ and $N(E) \propto E^{1.5}$, for sample sizes of $n = 100$ (top row), $n = 10^3$ (middle row), and $n = 10^4$ (bottom row). Note the truncation bias for a threshold at $P \geq 0.5$, which causes a lower rollover in the frequency distributions. The parameter correlations $E \propto P^{\beta}$ were fitted with a linear regression fit (left column) and the powerlaw slopes were fitted in the decreasing part on the right-hand side of the maximum of the distributions.

We bin the range of P and E each with 10 bins and determine the powerlaw slope with a linear regression for the bins right to the peak of the frequency distributions, in order to eliminate the rollover due to the sensitivity loss at low values, and find slopes of $\alpha_P = 2.04$ and $\alpha_E = 1.94$ (Fig. 7.4, top row), while we expect theoretically $\alpha_P = 2.0$ and $\alpha_E = 1.5$. Part of the discrepancy results from the small number statistics and the choice of bins in the linear regression fit, while part of the difference is caused by the generation of random numbers.

We repeat the same simulation for two larger sets of $n = 10^3$ (Fig. 7.4, middle row) and $n = 10^4$ events (Fig. 7.4, bottom row). We see that the correlation converges to a value of $\beta = 1.94$, which is still slightly different from the theoretical value of $\beta = 2.0$, either because of the truncation bias of the flux threshold or the random number generation. The powerlaw slopes of the distribution of peak energies P converges to $\alpha_P = 2.13$, and the slope of total energies E to $\alpha_E = 1.58$. The latter result is closer to the theoretical value of $\alpha_E = 1.5$, so increased statistics helps. The remaining difference in the powerlaw slope in the order of $\approx 5\%$ is caused by a combination of the truncation bias, the choice of fitted bins, and the random number generation. A bin-free powerlaw fitting procedure is described in Parnell and Jupp (2000), which may be preferable in some cases. A different option is also weighting of the bins by the number of events per bin. However, a best-fit in a log-log plot is generally achieved when equidistant bins on a logarithmic scale have equal weights, which is different from the weighting by number of events per bin. Whatever fitting strategy is considered best, we have always to keep in mind that the data often do not exactly correspond to a theoretical model, in which case a fitting parameter is in principle ill-defined. For instance, if the data obey an exponential distribution, a fit with a powerlaw model will yield a variable slope that starts from very flat at the lower end to very steep at the upper end of the distribution, so the powerlaw slope is ill-defined.

Let us also test the consistency between the linear regression fit of the correlation and the powerlaw fits of the frequency distributions. From the statistically largest sample we measure $\alpha_P = 2.13$ and $\alpha_E = 1.58$ (Fig. 7.4, bottom row), which provides a prediction of $\beta = (\alpha_P - 1)/(\alpha_E - 1) = (2.13 - 1)/(1.58 - 1) = 1.948$ that indeed agrees with the linear regression fit of the correlation plot, $\beta = 1.94$, with high accuracy ($\lesssim 0.4\%$), confirming the robustness of the fitting procedures used in this case. The discrepancy to the theoretical values of $\beta = 2$ seems to be caused by the particular random number representation used in the Monte-Carlo simulation for this case.

7.2 Frequency Distributions in Magnetospheric Physics

Let us now turn to observed occurrence frequency distributions of SOC events, starting from the magnetosphere. Some examples of frequency distributions of area sizes and dissipated power of magnetospheric substorm events are shown in Fig. 1.10, and lifetime distributions of substorm-related events are shown in Fig. 7.5, which are listed in Table 7.1.

Table 7.1 Frequency distributions measured in magnetospheric physics. References: 1, Lui et al. (2000); 2, Angelopoulos et al. (1999); 3, Takalo (1993); 4, Takalo et al. (1999a); 5, Freeman et al. (2000b); 6, Chapman and Watkins (2001), 7, Crosby et al. (2005).

Phenomenon	Parameter	Powerlaw slope α	Reference
Substorms (active)	area size	1.21 ± 0.08	1
	power	1.05 ± 0.08	1
Substorms (quiet)	area size	1.16 ± 0.03	1
	power	1.00 ± 0.02	1
Substorms	flow burst durations	1.59 ± 0.07	2
AE index	lifetimes	1.24	3,4
AU index	burst lifetimes	1.3	5,6
Outer radiation belt	electron counts	1.5–2.1	7

Frequency distributions of substorms have been obtained by measuring the projected area (in units of square kilometers) of auroral blobs with the POLAR spacecraft (Fig. 1.9), as well as by measuring the dissipated power (in units of watts) with the Ultraviolet Imager UVI (Lui et al. 2000). As a first-order approximation we can consider both the area or the dissipated power as a measure of the total energy E of substorm events. The frequency distributions (Fig. 1.10) show powerlaw distributions during quiet time intervals with slopes of $\alpha_E \approx 1.00$–1.16, and a similar slope of $\alpha_E \approx 1.05$–1.21 during substorm active time intervals, although there is in addition a Gaussian hump at the upper end of the distribution, which has been interpreted as a finite-size effect (Chapman et al. 1998), as simulated with a numerical model (Fig. 2.13; Section 2.5.1). Count rates of electrons accumulated by microsatellites during each crossing of the Earth's outer radiation belt revealed also powerlaw distributions, with slopes of $\alpha_E \approx 1.5$–2.1 (Crosby et al. 2005). It would be interesting to compare the same dataset $N(E)$ with the frequency distributions $N(P)$ of peak energies and $N(T)$ of burst durations, in order to test SOC models. However, we find related frequency distributions of burst durations measured from the AE index (Fig. 7.5 top; Takalo 1993) and AU index (Fig. 7.5 bottom; Freeman et al. 2000b), as well as from the durations of bursty bulk flow bursts in the magnetotail plasma sheet (Fig. 7.5 middle; Angelopoulos et al. 1999), which all exhibit powerlaw-like distributions in the range of $\alpha_T \approx 1.2$–1.6 (Table 7.1).

If we interpret these magnetospheric substorms as SOC events and apply our standard model of SOC avalanches with exponential growth and linear decay (Section 3.1), we expect the following relations for powerlaw slopes as summarized in Eq. (3.1.28),

$$\begin{aligned} \alpha_P &= 1 + \tau_G/t_S \\ \alpha_T &= \alpha_P \\ \alpha_E &= (\alpha_P + 1)/2 \end{aligned} \quad . \tag{7.2.1}$$

The fact that powerlaw slopes of $\alpha_T \approx 1.24$–1.6 are measured in substorms, would indicate that the ratio of the exponential growth rate τ_G to the mean saturation time t_S has a relatively low value of $(\tau_G/t_S) = \alpha_P - 1 = \alpha_T - 1 = 0.2$–$0.6$, which implies that the responsible instability saturates after $(t_S/\tau_G) = 1.6$–5.0 growth times. This implies relatively large amplification factors of $\exp(t_S/\tau_G) \approx 5$–$150$, which could be verified from the exponential growth during the rise time of substorm events. Such high amplification factors require coherent growth without a competing damping mechanism or collisional interactions, a characteristic that could constrain possible physical mechanisms responsible for geomagnetic substorms. Another prediction from our standard model is the powerlaw slope of the frequency distribution of energies, which based on $\alpha_T \approx 1.2$–1.6 would fall according to Eq. (7.2.1) in the range of $\alpha_E = (\alpha_P + 1)/2 = (\alpha_T + 1)/2 \approx 1.1$–$1.3$, which indeed is close to the values that Lui et al. (2000) observed during active time intervals of substorms, i.e., $\alpha_E \approx 1.21 \pm 0.08$ for the areas of substorms, and $\alpha_E \approx 1.05 \pm 0.08$ for the dissipated power, respectively.

In summary, we can conclude that the observed frequency distribution of magnetospheric substorms exhibit powerlaw-like functions for energy E and time duration T parameters, which are consistent with the statistics of an exponentially growing insta-

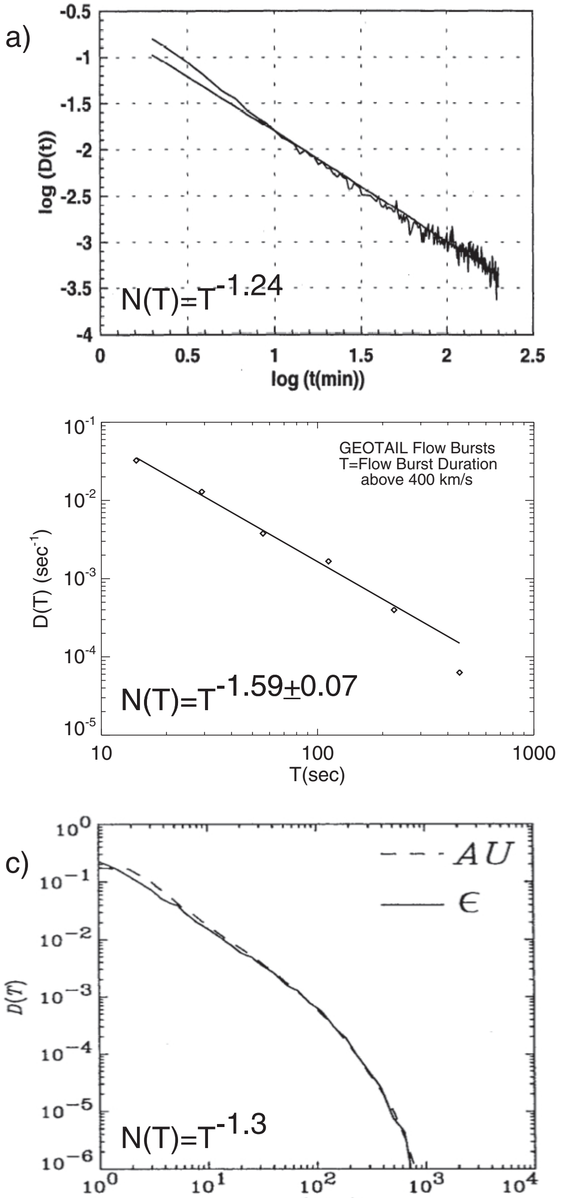

Fig. 7.5 *Top (a):* Lifetime distribution of magnetospheric disturbances as measured from the AE index of the years 1978–1985 (Takalo 1993; 1999a); *Middle (b):* Probability density distribution $N(T)$ of continuous flow magnitude durations T measured with the *GEOTAIL* satellite when it encountered the magnetotail plasma sheet during the period of Jan 1996–Oct 1998 (Angelopoulos et al. 1999); *Bottom (c):* Probability density function $N(T)$ of lifetimes of bursty bulk flow events in substorms measured of AU (Jan 1978–Jun 1988) and ε calculated from WIND SWE and MFI data for 1984–1987 (Freeman et al. 2000b; Chapman and Watkins 2001).

bility that has saturation times in the order of $t_S/\tau_G \approx 1.6$–5.0, amplification factors of $\exp(t_S/\tau_G) \approx 5$–$150$, leading to powerlaw slopes of $\alpha_T = \alpha_P \approx 1.2$–$1.6$ for durations, and $\alpha_E = 1.1$–1.3 for energies of substorm events.

7.3 Frequency Distributions in Solar Physics

Solar flares are probably the best-studied datasets regarding SOC statistics in astrophysics. Solar flares are catastrophic events in the solar corona, most likely caused by a magnetic instability that triggers a magnetic reconnection process, producing emission in almost all wavelengths, such as in gamma rays, hard X-rays, soft X-rays, extreme ultraviolet (EUV), Hα emission, radio wavelengths, and sometimes even in white light. Since the emission mechanisms are all different in each wavelength, such as nonthermal bremsstrahlung (in hard X-rays), thermal bremsstrahlung (in soft X-rays and EUV), gyrosynchrotron emission (in microwaves), plasma emission (in metric and decimetric waves), etc., we expect that the calculation of energies contained in each event strongly depends on the emission mechanism, and thus on the wavelength. It is therefore advisable to investigate the statistics of SOC events in each wavelength domain separately. The most unambiguous SOC parameters to report are the peak flux P, the total flux or fluence E, defined as the time-integrated flux over the entire event, and the total time duration T of the event. Conversions of fluxes and fluences into energy release rates and total energies require physical models, which will be discussed in Chapter 9.

7.3.1 Solar Flare Hard X-rays

Hard X-ray emission in solar flares mostly results from thick-target bremsstrahlung of nonthermal particles accelerated in the corona that precipitate into the dense chromosphere. Thus, the hard X-ray flux is the most direct measure of the energy release rate, and thus is expected to characterize the energy of SOC events in a most uncontaminated way, while emission in other wavelengths exhibit a more convolved evolution of secondary emission processes.

One of the earliest reports of a frequency distribution of solar hard X-ray flare fluxes was made by Datlowe et al. (1974), who published the cumulative frequency distribution of 123 flare events detected in the 20–30 keV energy range above a threshold of $\gtrsim 0.1$ photons (cm^{-2} s^{-1} keV^{-1}) with the OSO-7 spacecraft during 10 Oct 1971–6 June 1972, finding a powerlaw slope of $\beta_P \approx 0.8$. For compatibility we list only powerlaw slopes of differential frequency distributions in Table 7.2, and use the conversion $\alpha = \beta + 1$ if needed (but see bias described in Section 7.1.2). We list also the logarithmic ranges of the x-axis over which the powerlaw fit was obtained, e.g., ${}^{10}\log(P_{max}/P_{min}) = {}^{10}\log(30/0.3) \approx 2$ in the case of Datlowe et al. (1974), which is a good indicator of the accuracy of the powerlaw slope fit.

A sample of 25 microflares of smaller size were detected at 20 keV with a balloon-borne instrumentation of *University of California Berkeley (UCB)* during 141 minutes of

Table 7.2 Frequency distributions measured from solar flares in hard X-rays and gamma-rays. References: 1, Datlowe et al. (1974); 2, Lin et al. (1984); 3, Dennis (1985); 4, Schwartz et al. (1992); 5, Crosby et al. (1993); 6, Biesecker et al. (1993); 7, Biesecker et al. (1994); 8, Crosby (1996); 9, Lu et al. (1993); 10, Lee et al. (1993); 11, Bromund et al. (1995); 12, Perez-Enriquez and Miroshnichenko (1999); 13, Georgoulis et al. (2001); 14, Su et al. (2006); 15, Christe et al. (2008); 16, Lin et al. (2001); 17, Tranquille et al. (2009).

Powerlaw slope of peak flux α_P	Powerlaw slope of total fluence α_E	Powerlaw slope of durations α_T	log range	Instrument	Reference
1.8			2	OSO-7	1
2.0			1	UCB	2
1.8			4	HXRBS	3
1.73 ± 0.01			3.5	HXRBS	4
1.73 ± 0.01	1.53 ± 0.02	2.17 ± 0.05	3.5	HXRBS	5
1.61 ± 0.03			3.5	BATSE	4
1.75 ± 0.02			4	BATSE	6
1.68 ± 0.02			3.5	BATSE	7
1.59 ± 0.02		2.28 ± 0.08	3	WATCH	8
1.86	1.51	1.88	3	ISEE-3	9
1.75	1.62	2.73	2.5	ISEE-3	10
1.86 ± 0.01	1.74 ± 0.04	2.40 ± 0.04	3.5	ISEE-3	11
1.80 ± 0.01	1.39 ± 0.01		1	PHEBUS	12
1.59 ± 0.05	1.39 ± 0.02	1.09-1.15	2	WATCH	13
1.80 ± 0.02		3.6 [0.9]	3.5	RHESSI	14
1.58 ± 0.02	1.7 ± 0.1	2.2 ± 0.2	2	RHESSI	15
1.6			3	RHESSI	16
1.61 ± 0.04			2	ULYSSES	17

observations on 1980 June 27, yielding a powerlaw distribution with a slope of $\beta \approx 1$ (Lin et al. 1984).

A much larger amount of statistics was obtained with the *Hard X-Ray Burst Spectrometer (HXRBS)* onboard the *Solar Maximum Mission (SMM)* spacecraft, which recorded 6,775 flare events during the 1980–1985 period, exhibiting a powerlaw distribution of peak count rates with a slope of $\alpha_P = 1.8$ over four orders of magnitude (Dennis 1985), see Fig. 1.13.

A next mission with hard X-ray detector capabilities was the *Compton Gamma Ray Observatory (CGRO)*. Although it was designed to detect gamma-ray flashes from astrophysical objects, it detected also solar flares systematically during the period of 1991–2000. Using the *Burst And Source Transient Experiment (BATSE)*, statistics of flares with energies >25 keV was sampled and more detailed powerlaw distributions of peak fluxes were reported with values of $\alpha_P = 1.61 \pm 0.03$ (Schwartz et al. 1992), $\alpha_P = 1.75 \pm 0.02$ (Biesecker et al. 1993). and $\alpha_P = 1.68 \pm 0.02$ (Biesecker et al. 1994) for BATSE. Biesecker et al. (1994) noticed slight differences of the powerlaw slope during low activity ($\alpha_P = 1.71 \pm 0.04$) and high activity periods ($\alpha_P = 1.68 \pm 0.02$), which appear not to be significant.

A systematic study of flares observed with HXRBS over the entire mission duration of 1980–1989 was conducted by Crosby et al. (1993), measuring peak count rates P_{cts} (cts

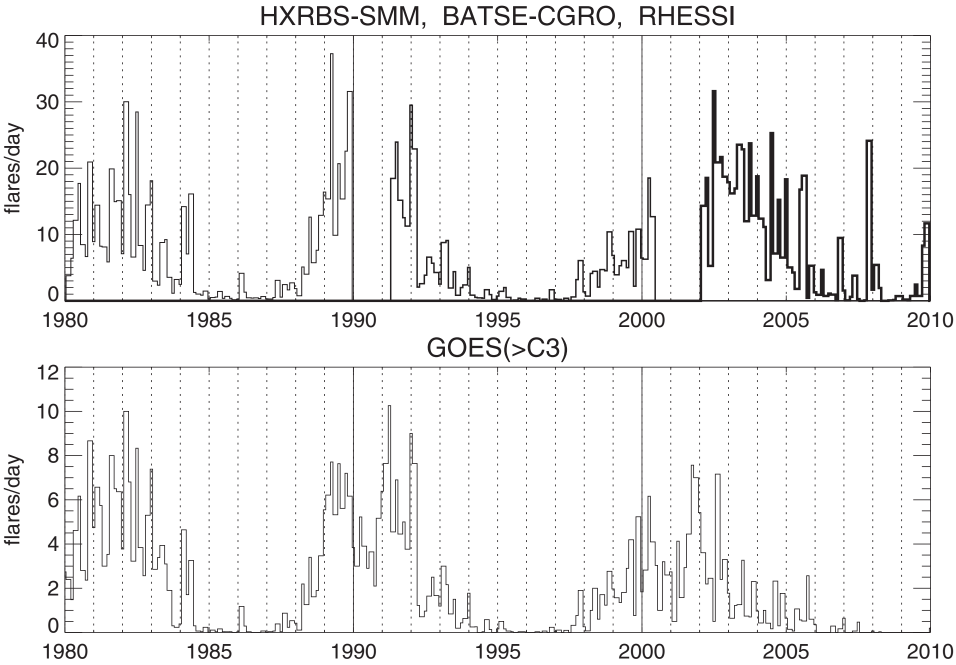

Fig. 7.6 *Top:* Monthly averages of solar flare rates observed during the last three solar cycles in hard X-rays with HXRBS/SMM (1980–1989), BATSE/CGRO (1991–2000), and RHESSI (2002–2010), corrected for the duty cycles of the instruments. *Bottom:* Monthly averages of the solar flare rate observed in soft X-rays with GOES, including events above the C3-class level.

s^{-1}), converted into photon fluxes P_{ph} (photons $\mathrm{cm}^{-2}\ \mathrm{s}^{-1}\ \mathrm{keV}^{-1}$) at energies >25 keV, peak HXR spectrum-integrated fluxes P_X (photons $\mathrm{cm}^{-2}\ \mathrm{s}^{-1}$), peak electron fluxes P_e (ergs s^{-1}), flare durations T, and time-integrated total energies in electrons E_e (ergs), for four different time intervals of the solar cycle. The variability of the solar flare rate during the last three solar cycles can be seen in form of monthly averages in Fig. 7.6. In Table 7.2 we list the values for the time range of 1980–1982, which covers the solar maximum and has the largest statistics. The values of the powerlaw slopes change only by $\lesssim 2\%$ during the solar minimum. The multi-parameter statistics of P, E, and T allowed also to derive the following parameter correlations (see Section 7.1.6),

$$\begin{aligned} T &\propto P_{ph}^{0.43[0.41]} \\ P_X &\propto P_{ph}^{0.95[1.06]} \\ P_e &\propto P_{ph}^{1.02[0.93]} \\ E_e &\propto P_{ph}^{1.21[1.25]} \end{aligned} , \qquad (7.3.1)$$

where the powerlaw index is derived by linear regression between the parameters, as well as from the slopes of each frequency distribution with Eq. (7.1.42) (indicated in brackets

[...]). The peak parameters seem to be all close to proportional to each other, i.e., $P_X \propto P_e \propto P_{ph}$, so it does not matter much which one is used to characterize the peak energies of SOC avalanches. It is interesting to compare these correlation coefficients with our standard SOC model (Section 3.1), which predicts $\alpha_P = \alpha_T$ and $\alpha_E = (\alpha_P + 1)/2$, and the correlations $T \propto P^1$ and $E \propto P^2$. We have to investigate observations in other wavelengths and explore whether different definitions of event durations and energies can explain the discrepancy of observed correlations (Eq. 7.3.1) to the theoretical model.

From the *Wide Angle Telescope for Cosmic Hard X-Rays (WATCH)* onboard the Russian satellite *Granat*, a sample of 1,546 flare events was observed at energies of 10–30 keV or 14–40 keV during 1990–1992, yielding similar powerlaw slopes for peak count rates, $\alpha_P = 1.59 \pm 0.02$, and flare durations $\alpha_T = 2.28 \pm 0.08$ as reported before (Crosby 1996; Crosby et al. 1998). However, it was noted that the frequency distribution of flare durations exhibits a gradual rollover for short flare durations, approaching a slope of $\alpha_T \approx 1$, so it cannot be fitted with a single powerlaw distribution over the entire range of flare durations. From the *PHEBUS* instrument on *Granat*, which is sensitive to gamma-ray energies, Perez-Enriquez and Miroshnichenko (1999) analyzed 110 high-energy solar flares observed in the energy range of 100 keV–100 MeV and found the following powerlaw slopes: $\alpha_P = 1.80 \pm 0.01$ for (bremsstrahlung) hard X-ray fluxes at >100 keV, $\alpha_P = 1.38 \pm 0.01$ for photon energies at 0.075–124 MeV, $\alpha_P = 1.39 \pm 0.01$ for bremsstrahlung at 300–850 keV, $\alpha_E = 1.50 \pm 0.03$ for the 511 keV electron-positron annihilation line fluence, $\alpha_E = 1.39 \pm 0.02$ for the 2.223 MeV neutron capture line fluence, and $\alpha_E = 1.31 \pm 0.01$ for the 1–10 MeV gamma-ray line fluence. We have to be aware that this selection of high-energy (gamma-ray) flares is not representative for all hard X-ray flares, and thus has a biased distribution towards the largest events, which explains that most frequency distributions in gamma rays have a flatter slope than in hard X-rays. The flatter slope corresponds according to our standard model ($\alpha_P = 1 + t_S/\tau_G$) also to events with higher exponential growth factors $\exp(\tau/t_S)$, which is certainly expected for gamma ray-producing flare events.

Using data from a >25 keV hard X-ray detector onboard the ISEE-3/ICE spacecraft during 24 Aug 1978 and 11 Jul 1986, Lu et al. (1993) determined the frequency distributions of the peak luminosity P (erg s^{-1}), the energy E (erg), and flare duration T (s) and found that the measured distributions could be best fitted with a cellular automaton model that produced powerlaw slopes of $\alpha_P = 1.86$, $\alpha_E = 1.51$, and $\alpha_T = 1.88$. The fits of the distributions included an exponential rollover at the upper end, which explains that they inferred a less steep slope for durations than previously reported. Interestingly, these values agree much more closely with our standard model, which predicts for $\alpha_P = 1.86$ the slopes $\alpha_T = \alpha_P = 1.86$ and $\alpha_E = (\alpha_P + 1)/2 = 1.43$. This tells us that the rollovers at the lower and upper end of the distributions have to be included in the model fits in order to obtain proper powerlaw slopes. Lee et al. (1993) analyzed the same data and determined the correlations and frequency distribution powerlaw slopes with special care of truncation biases and obtained similar values for ISEE-3 ($\alpha_P = 1.75$, $\alpha_E = 1.62$, $\alpha_T = 2.73$) as Crosby et al. (1993) for HXRBS. A third study was done with the same data (Bromund et al. 1995), where the energy spectrum was also calculated to determine different energy parameters, similar to the study of Crosby et al. (1993), finding the following powerlaw slopes: $\alpha_P = 1.86, ..., 2.00$ for the peak photon flux P_{ph} (photons cm^{-2} s^{-1}),

$\alpha_P = 1.92, ..., 2.07$ for the peak electron power P_e (erg s^{-1}), $\alpha_E = 1.67, ..., 1.74$ for the total electron energy E_e (erg), and $\alpha_T = 2.40, ..., 2.94$ for the total duration T (s), where the range of powerlaw slopes results from the choice of the fitting range. The flare duration T was defined at a level of 1/e times the peak count rate. The following parameter correlations were found,

$$\begin{array}{l} P \propto T^{0.75[1.52]} \\ E \propto T^{1.60[2.08]} \\ E \propto P^{1.35[1.36]} \end{array} , \qquad (7.3.2)$$

where the powerlaw index is derived by orthogonal linear regression fits, as well as from the slopes of the frequency distributions Eq. (7.1.42) (indicated in brackets [...]). Interestingly, the first two correlations are consistent with our standard model, which predicts $P \propto T^1$ and $E \propto T^2$, within the uncertainty of the two methods.

From the latest solar mission with hard X-ray capabilities, the *Ramaty High-Energy Solar Spectroscopic Imager (RHESSI)* spacecraft, frequency distributions were determined in the 12–25 keV energy band from 2002–2005 (Su et al. 2006), finding powerlaw slopes of $\alpha_P = 1.80 \pm 0.02$ for the peak fluxes, and a broken powerlaw $\alpha_T = 0.9$–3.6 for the flare duration, similar to previous findings (e.g., Crosby et al. 1998). Christe et al. (2008) conducted a search of microflares and identified a total of ≈25,000 events observed with RHESSI during 2002–2007 and investigated the frequency distributions at lower energies, finding powerlaw slopes of $\alpha_P = 1.50 \pm 0.03$ for 3–6 keV peak count rates P_{ph} (cts s^{-1}), $\alpha_P = 1.51 \pm 0.03$ for 6–12 keV peak count rates, and $\alpha_P = 1.58 \pm 0.02$ for 12–25 keV peak count rates. Converting the peak count rates P into total energy fluxes by integrating their energy spectra, Christe et al. (2008) find an energy distribution with a powerlaw slope of $\alpha_E = 1.7 \pm 0.1$, with an average energy deposition rate of $\lesssim 10^{26}$ erg s^{-1}. It is interesting that this microflare statistics is fairly consistent with overall flare statistics, even if it represents only a subset in the lowest energy range.

Flare statistics was also gathered from the *Solar X-ray/Cosmic Gamma-Ray Burst Experiment (GRB)* onboard the *Ulysses* spacecraft (Tranquille et al. 2009), finding similar results for >25 keV events, i.e., a powerlaw slope of $\alpha_P = 1.61 \pm 0.04$ for the peak count rate, which steepens to $\alpha_P = 1.75 \pm 0.08$ if the largest events with pulse pile-up are excluded.

A specialized study investigated also how the frequency distribution of hard X-ray peak fluxes depends on the associated size of the active region and found evidence for an upper cutoff due to a finite size limit (Kucera et al. 1997). Besides flare events *per se*, one can also consider substructures in flares and conduct SOC statistics. If the substructures are self-similar to the overall structure, one would expect similar powerlaw slopes of the frequency distributions, which was indeed found to be the case for hard X-ray subpulses (Aschwanden et al. 1995), although subpulses from a single flare can have exponential distributions, while their superposition from many flares converges towards powerlaw distributions (Aschwanden et al. 1998b). These subpulses have typical time scales of $T_{sub} = 0.5$–1.5 s (Aschwanden et al. 1995) and $T_{sub} = 1.9 \pm 0.5$ s (Qiu and Wang 2006).

A summary plot of frequency distributions of the peak count rate P is shown for the three instruments HXRBS/SMM, BATSE/CGRO, and RHESSI in Fig. 7.7, yielding an average powerlaw slope of $\alpha_P = 1.75 \pm 0.05$. The corresponding frequency distributions

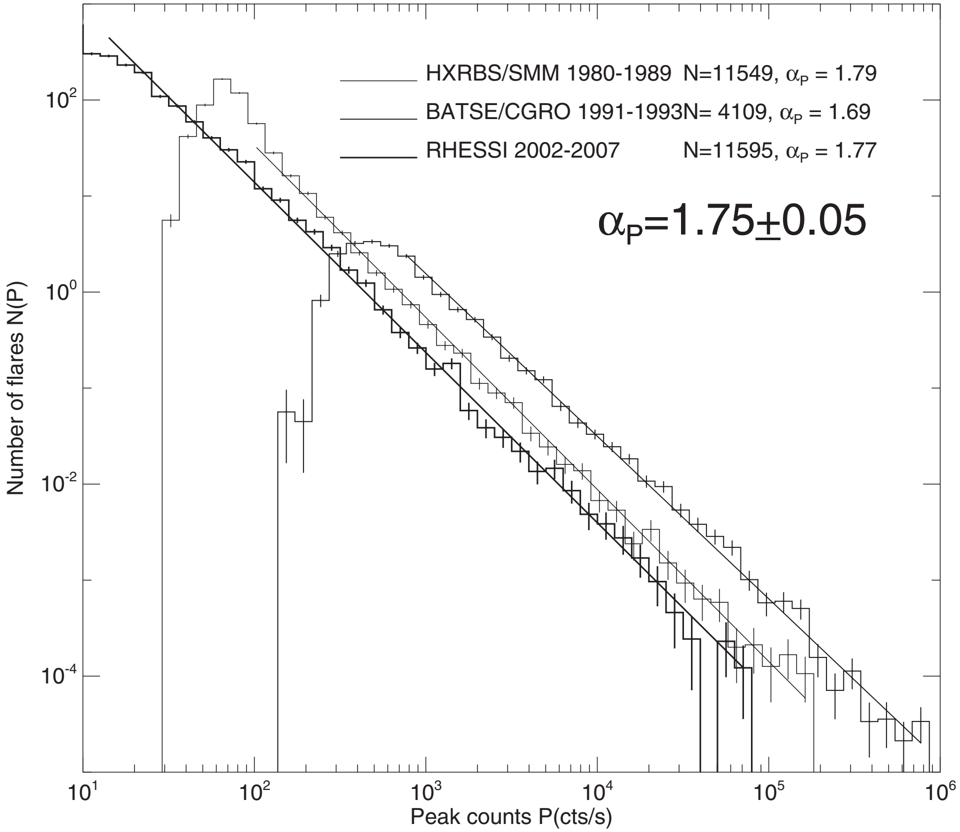

Fig. 7.7 Occurrence frequency distributions of hard X-ray peak count rates $P(cts/s)$ observed with HXRBS/SMM (1980–1989), BATSE (1991–1993), and RHESSI (2002–2007), with powerlaw fits. Note that BATSE/CGRO has larger detector areas, and thus records higher count rates. RHESSI flares were detected at energies of ≥12 keV, while HXRBS and BATSE flares were detected at energies of ≥25 keV. The average slope value is $\alpha_P = 1.75 \pm 0.05$.

of total counts or fluences are shown in Fig. 7.8, which have an average powerlaw slope of $\alpha_E = 1.61 \pm 0.04$. The distributions of flare durations are shown in Fig. 7.9, which exhibit an average of $\alpha_T = 2.08 \pm 0.10$, with a tendency toward a rollover at the low end. Thus, our best values are,

$$\begin{aligned} N(P) &\propto P^{-\alpha_P} \qquad \alpha_P = 1.75 \pm 0.05 \\ N(E) &\propto E^{-\alpha_E} \qquad \alpha_E = 1.61 \pm 0.04 \\ N(T) &\propto T^{-\alpha_T} \qquad \alpha_T = 2.08 \pm 0.10 \end{aligned} \tag{7.3.3}$$

It is interesting to note that these best observational values closely correspond to the numerically simulated values in the cellular automaton model of Lu and Hamilton (1991), see Eq. (2.6.15), i.e., $N(P) \propto P^{-1.67\pm0.04}$, $N(E) \propto E^{-1.53\pm0.02}$, and $N(T) \propto T^{-2.17\pm0.05}$. Using the averaged values of Eq. (7.3.3), we expect the following correlations between

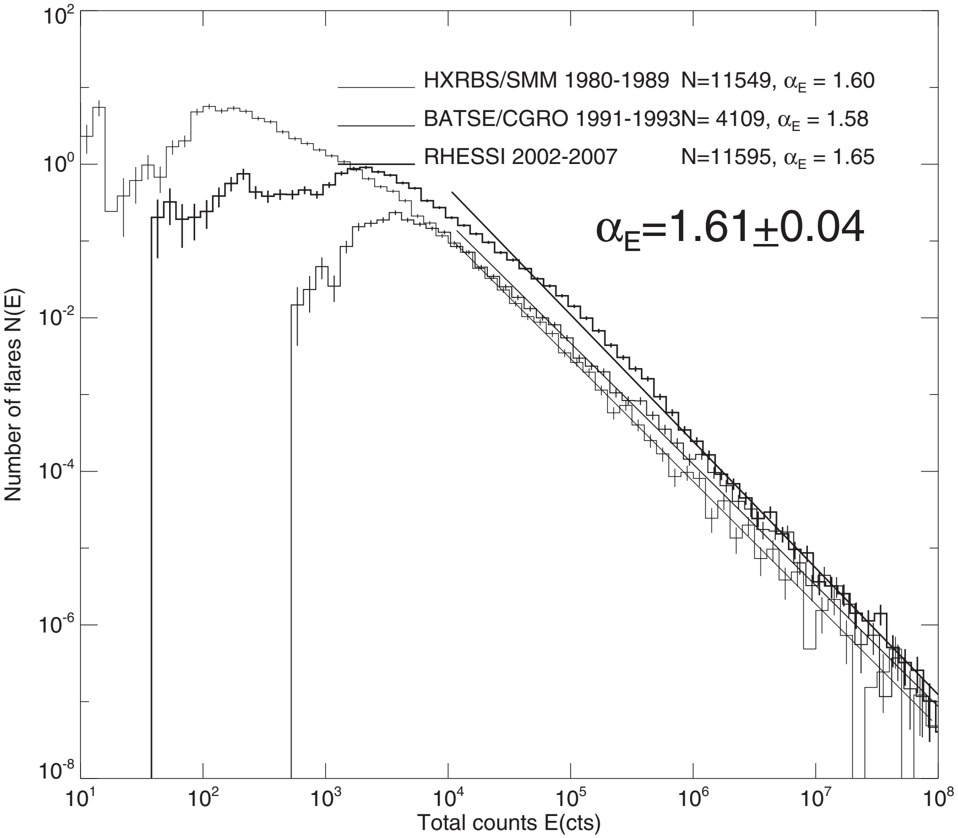

Fig. 7.8 Occurrence frequency distributions of hard X-ray total counts or fluence $E(cts)$ observed with HXRBS/SMM (1980–1989), BATSE (1991–1993), and RHESSI (2002–2007), with powerlaw fits. The average slope value is $\alpha_E = 1.61 \pm 0.04$.

these three parameters (using Eq. 7.1.42),

$$\begin{array}{ll} E \propto P^{\beta} & \beta = (\alpha_P - 1)/(\alpha_E - 1) = (1.75 - 1)/(1.61 - 1) = 1.23 \pm 0.09 \\ T \propto P^{\beta} & \beta = (\alpha_P - 1)/(\alpha_T - 1) = (1.75 - 1)/(2.08 - 1) = 0.70 \pm 0.07 \\ E \propto T^{\beta} & \beta = (\alpha_T - 1)/(\alpha_E - 1) = (2.08 - 1)/(1.61 - 1) = 1.77 \pm 0.16 \end{array} \quad (7.3.4)$$

In Fig. 7.10 we show the actual correlation plots between the parameters and determine linear regression fits, which give a comparable result, with $E \propto P^{1.26\pm0.04}$ and $T \propto P^{0.44\pm0.04}$. The latter correlation, of course, cannot be determined accurately from linear regression fits due to the large scatter in flare duration values. Thus, we have obtained representative values of the powerlaw slopes α and correlation coefficients β for solar flare hard X-ray parameters, averaged from three major missions over the last 30 years and three solar cycles, which can serve as reference for other wavelengths.

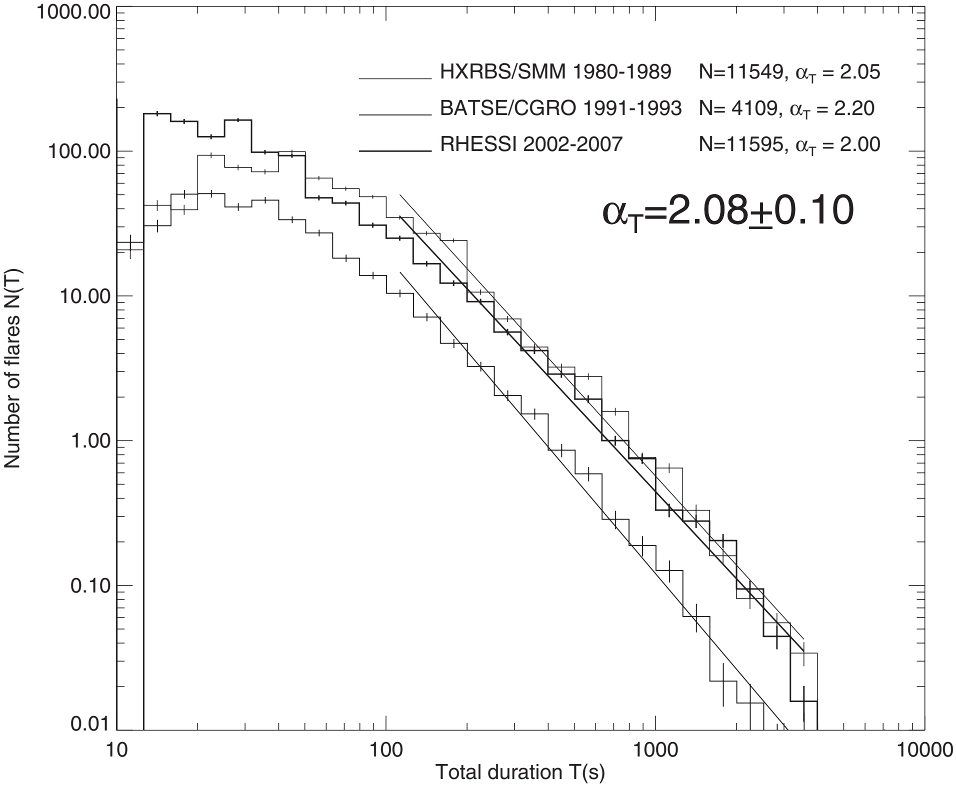

Fig. 7.9 Occurrence frequency distributions of hard X-ray flare durations $T(s)$ observed with HXRBS/SMM (1980-1989), BATSE (1991-1993), and RHESSI (2002-2007), with powerlaw fits. The flare durations for RHESSI were estimated from the time difference between the start and peak time, because RHESSI flare durations were determined at a lower energy of 12 keV (compared with 25 keV for HXRBS and BATSE), where thermal emission can dominate in large flares, causing a flatter powerlaw slope ($\alpha_T \approx 1.4$). The average slope value is $\alpha_T = 2.08 \pm 0.10$.

7.3.2 Solar Flare Soft X-rays

Soft X-ray emission in solar flares mostly originates from free-free bremsstrahlung emission of heated flare plasma, which typically reaches temperatures of $T \approx$ 10–35 MK. A pragmatic relationship between soft and hard X-ray emission of flare plasmas is characterized with the so-called *Neupert effect* (e.g., Dennis and Zarro 1993), which essentially states that the time profile of hard X-ray emission corresponds to the heating rate produced by nonthermal particles bombarding the chromosphere, while soft X-ray emission represents the chromospheric response of flare plasma heating. This model is called the *chromospheric evaporation* scenario. The thermal energy of the heated plasma and thus the time profile of emitted soft X-ray emission consequently approximately follows the time integral of the hard X-rays, until cooling by thermal conduction and radiative loss

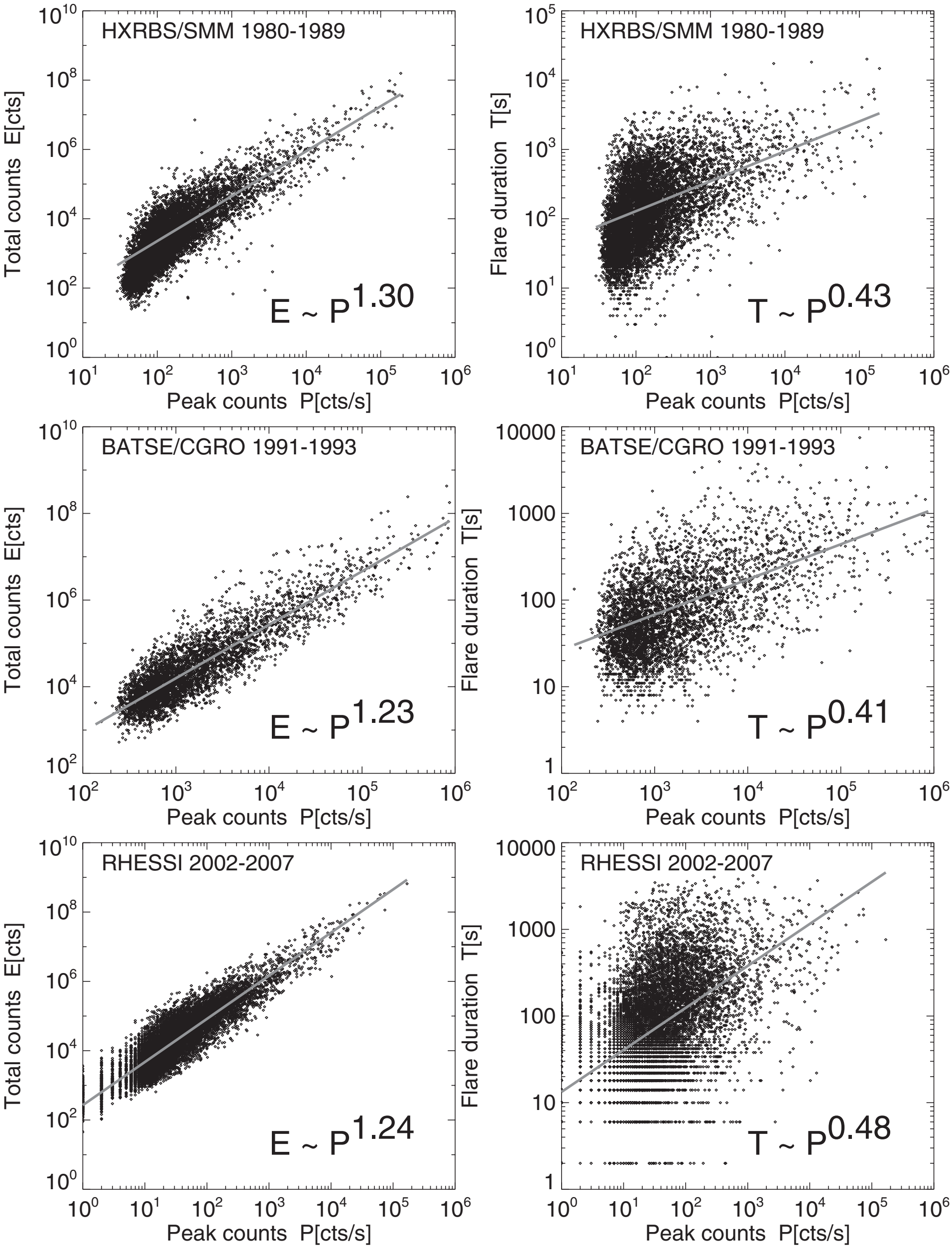

Fig. 7.10 Scatterplots between for the total counts $E(P)$ (left panels) or flare duration $T(P)$ (right panels) versus the peak count rate P for solar flares with HXRBS/SMM (1980–1989) (top), BATSE/CGRO (1991–1993) (middle), and RHESSI (2002–2007) (bottom). Linear regression fits are applied to all datapoints above a threshold of five times the minimum value in each parameter. All data are subject to a flux threshold P_{min}, which causes a truncation at $P \leq P_{min}$, but does not affect linear regression fits of the form $y(x)$.

overcomes the heating rate in the late flare phase. From this scenario we expect the relationship,

$$F^{SXR}(t) \approx \int_0^t F^{HXR}(t')\, dt' \ . \tag{7.3.5}$$

If we define a SOC event by the energy release as observed in hard X-rays, characterized with a flare start time t_s, end time t_e, total duration $T = (t_e - t_s)$, peak energy flux $P^{HXR} = F(t = t_p)$, and total flux or fluence $E = \int_{t_s}^{t_e} F^{HXR}(t)\, dt$, then the peak time t_p^{HXR} of the hard X-rays corresponds to the inflection point with the steepest rise in the soft X-ray time profile, while the end time t_e^{HXR} corresponds to the peak time in soft X-rays. In order to obtain a consistent flare duration T in the two wavelengths, we have therefore to define,

$$T = (t_e^{HXR} - t_s^{HXR}) = (t_p^{SXR} - t_s^{HXR}) \ , \tag{7.3.6}$$

and to calculate the time derivative of the soft X-ray light curve $F^{SXR}(t)$ according to the Neupert effect (Eq. 7.3.5),

$$F^{proxi}(t) = \frac{dF^{SXR}(t)}{dt} \ , \tag{7.3.7}$$

to obtain a proxy $F^{proxi}(t)$ for the hard X-ray-like flare light curve where we can measure the peak energies P and total energies E. If we do not correct for this Neupert effect, we expect some significantly different frequency distributions and correlation parameters for flare event statistics in soft X-rays and hard X-rays.

First frequency distributions of flare peak fluxes in soft X-rays were reported from OSO-3 observations in the energy range of 7.7–12.5 keV (Hudson et al. 1969), where a cumulative distribution with a powerlaw tail with a slope of $\beta \approx 0.8$ was found which corresponds to a slope of $\alpha \approx \beta + 1 = 1.8$ for the differential frequency distribution. Further data in the 2–12 Å range (1–6 keV) with *Explorer 33* and *Explorer 35* satellites yielded solar flare statistics for ≈3,000 events during July 1966 and September 1968, from which powerlaw distributions of the peak flux ($\alpha_P = 1.75$) and the fluence ($\alpha_E = 1.44$) were reported (Drake 1971).

The Yohkoh mission (1991–2002) provided imaging observations of solar flares with the *Soft X-ray Telescope (SXT)* at temperatures of $T > 1.5$ MK (>0.13 keV). Shimizu (1995) analyzed small active region transient brightenings (small flares) during August 1992 and inferred from a sample of some 5,000 events in a single active region frequency distributions of soft X-ray peak fluxes with powerlaw slopes in the range of $\alpha_P = 1.64$–1.89, depending on the spatial area used in the sampling. The thermal energy of these events were estimated in the range of $E \approx 10^{27}$–10^{29}, and the powerlaw slope for energies was calculated to $\alpha_E \approx 1.5$–1.6 (Shimizu 1995). A similar study was performed by Shimojo and Shibata (1999), who analyzed 92 microflares during the lifetime of a single bright point (i.e., a miniature active region) and found a power law slope of $\alpha_P = 1.7 \pm 0.4$ for the soft X-ray peak flux.

The difference between frequency distributions sampled in hard X-rays and soft X-rays was modeled by Lee et al. (1995). For this purpose, flare statistics in hard X-rays (HXRBS, ISEE-3) and in soft X-rays (SMM/BCS, GOES) were reanalyzed (Fig. 7.11), but similar powerlaw slopes were found for the two wavelength ranges, which could only

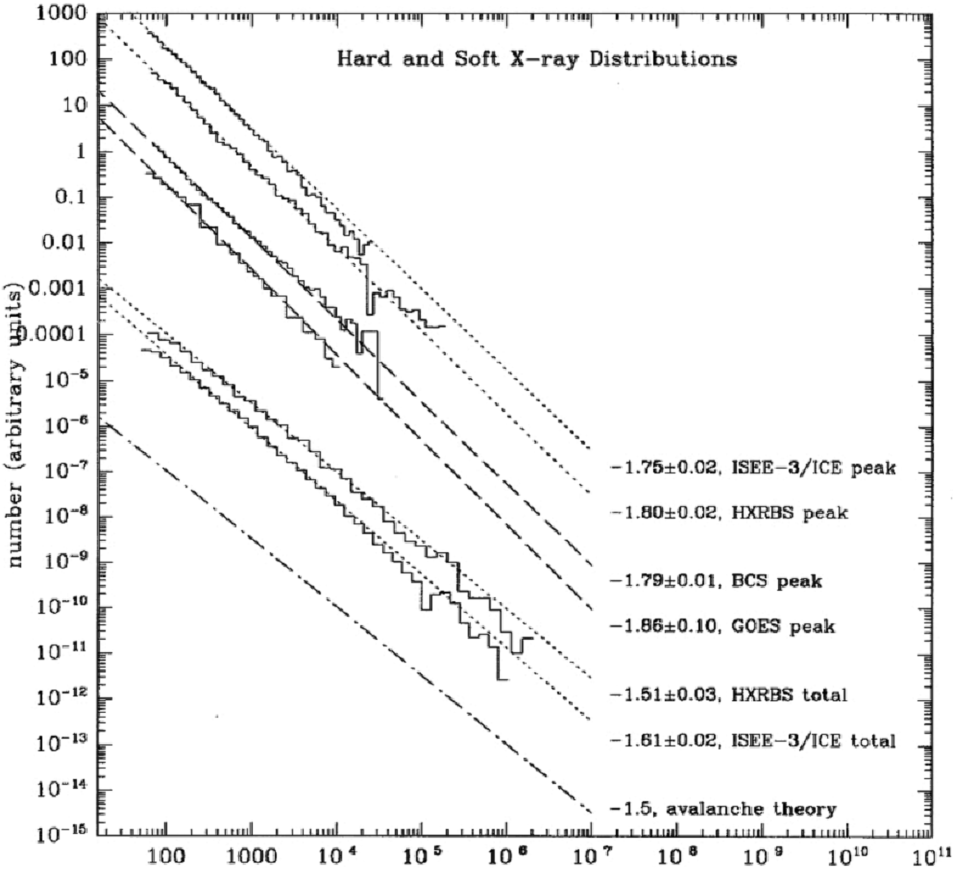

Fig. 7.11 Frequency distributions of peak fluxes (*peak*) and total fluxes (*total*) for soft X-ray (SMM/BCS, GOES) and hard X-ray (ISEE-3, HXRBS) flare events. The occurrence rates are arbitrarily scaled. Note the similar slopes in the two wavelength ranges (Lee et al. 1995; reproduced by permission of the AAS).

be reconciled with the expected difference for the chromospheric evaporation scenario if one assumes a special scaling law between temperature and density, i.e., $n \propto T^{-4/5}$ (Lee et al. 1995).

Using soft X-ray light curves from the *Geostationary Operational Environmental Satellites (GOES)*, which observe the Sun uninterrupted thanks to multiple spacecraft, complete flare statistics can be gathered. Feldman et al. (1997) sampled during 1993–1995 some 1,000 flare events in the 1–8 Å (0.08–0.67 keV) range and inferred a soft X-ray peak flux distribution with a powerlaw slope of $\alpha_P = 1.88 \pm 0.21$. A more comprehensive study of 50,000 soft X-ray flares observed with GOES during 1976–2000 was performed by Veronig et al. (2002a,b). The obtained frequency distributions exhibit significantly steeper slopes than previously found, i.e., $\alpha_P = 2.11 \pm 0.13$ for the peak flux, $\alpha_E = 2.03 \pm 0.09$ for the fluence, and $\alpha_T = 2.93 \pm 0.12$ for durations. This discrepancy with previous statistics (Table 7.3) most likely arises from two facts: (1) no pre-event background flux was sub-

Table 7.3 Frequency distributions measured from solar flares in soft X-rays. References: 1, Hudson et al. (1969); 2, Drake et al. (1971); 3, Shimizu (1995); 4, Lee et al. (1995); 5, Feldman et al. (1997); 6, Shimojo and Shibata (1999); 7, Veronig et al. (2002d); 8, Veronig et al. (2002a); 9, Yashiro et al. (2006).

Powerlaw slope of peak flux α_P	Powerlaw slope of total fluence α_E	Powerlaw slope of durations α_T	log range	Instrument	Reference
1.8			1	OSO-3	1
1.75	1.44		2	Explorer	2
1.64–1.89	1.5–1.6		2	Yohkoh	3
1.79			2	SMM/BCS	4
1.86			2	GOES	4
1.88 ± 0.21			3	GOES	5
1.7 ± 0.4			2	Yohkoh	6
1.98	1.88		3	GOES	7,8
$2.11 \pm 0.13^*$	$2.03 \pm 0.09^*$	$2.93 \pm 0.12^*$	3	GOES	8
$2.16 \pm 0.03^*$	$2.01 \pm 0.03^*$	$2.87 \pm 0.09^*$	3	GOES	9

[*] No background subtracted.

tracted, which substantially overestimates the peak flux and fluence of small events, and thus causes a steeper powerlaw slope, and (2) the Neupert effect could explain some discrepancy with respect to hard X-rays, but there seems to be a small difference (Fig. 7.11) according to the comparison of Lee et al. (1995). The Neupert effect has been investigated by correlating the soft X-ray peak flux with the hard X-ray fluence (Eq. 7.3.5) and a strong correlation was found, but it is not strictly proportional as predicted by the chromospheric evaporation model (Veronig et al. 2002c). Similar values ($\alpha_P = 2.16 \pm 0.03$, $\alpha_E = 2.01 \pm 0.03$, and $\alpha_T = 2.87 \pm 0.09$ were inferred by Yashiro et al. (2006), but these distributions suffer from the same lack of background subtraction as the study of Veronig et al. (2002a,b), which causes a bias in overestimating the flux of weak events and thus steepens the powerlaw slopes. Moreover, the study of Yashiro et al. (2006) demonstrated that a subset of flare events with simultaneous CME events exhibited flatter powerlaw distributions, which is to be expected for any subset that contains preferentially larger events.

In summary, for flare events observed in soft X-rays, we can group the results into two categories: (1) event statistics with pre-flare background subtraction, and (2) without pre-flare background subtraction. From the compilation shown in Table 7.3 it is clear that each group produces quite consistent results among themselves, but they differ significantly due to the well-understood bias caused by neglecting background subtraction, especially in GOES data, where every light curve contains also the total soft X-ray emission from all other active regions on the solar surface besides a particular flare event (e.g., Bornmann 1990). Thus, ignoring the second group (Veronig et al. 2002a,b; Yashiro et al. 2006) in Table 7.3, we obtain the following averages from the first group,

$$\begin{array}{ll} N(P_{SXR}) \propto P^{-\alpha_P} & \alpha_P = 1.79 \pm 0.06 \\ N(E_{SXR}) \propto E^{-\alpha_E} & \alpha_E = 1.50 \pm 0.05 \end{array} , \qquad (7.3.8)$$

which is not much different from the values obtained in hard X-rays (Eq. 7.3.3), i.e., $N(P_{HXR}) \propto P^{-1.75\pm0.05}$ and $N(E_{HXR}) \propto E^{-1.61\pm0.04}$. Although we expect some difference due to the Neupert effect (Eq. 7.3.5) in the relationship between soft and hard X-rays, the dissimilarity in the frequency distributions of SOC parameters is apparently not large, either because the time profiles are close to self-similar, or because the soft X-ray light curves are not purely thermal emission, but contain also significant nonthermal emission as observed at higher energies in hard X-rays. Insight into these problems could be obtained by comparing SOC statistics obtained from GOES light curves directly versus event detection from the time derivative of the GOES light curve (Eq. 7.3.7).

7.3.3 Solar Flare Extreme Ultraviolet Emission

The evolution of a solar flare in different wavelengths can be best understood by their temperature dependence. In a large flare, plasma becomes heated to $T \approx 20$–35 MK, which produces bright emission in soft X-rays. Once the plasma cools down in the postflare phase, soft X-ray emission fades and extreme ultraviolet (EUV) emission becomes brighter, which is produced by free-free and bound-bound emission at temperatures of $T \approx 1$–2 MK. The systematic delay in the peak of the emission in different wavelengths can best be seen in multi-wavelength observations of a large flare, such as during the Bastille-Day (14 July 2000) flare shown in Fig. 7.12, where the timing of the peak emission in each wavelength is exactly ordered according to the temperature peak sensitivity of the different instruments. It peaks first in the *Yohkoh/HXT* 14–23 keV channel, which is sensitive to the highest temperature that occurred in the flare ($T \approx 35$ MK), then in the soft X-ray channels (GOES, Yohkoh/SXT), and finally in the EUV channels (TRACE 195, 171 Å), which are sensitive in the temperature range of $T \approx 1$–2 MK. Therefore, SOC statistics of flare events can in principle be performed in all these wavelengths, but the frequency distributions of peak flux (P), fluence (E), and durations (T) are not necessarily identical, unless the time profiles in the different wavelengths are self-similar. The comparison of the light curves from 7 different wavelength ranges shown in Fig. 7.12 suggests that the durations become systematically longer in wavelengths corresponding to cooler temperatures, which implies a temperature-dependent scaling between peak flux and duration, $P \propto T^{\beta}$. Large flares with total energies of $E \approx 10^{27}$–10^{32} erg are visible in hard X-rays and soft X-rays (Fig. 1.14), but tiny flares with energies of $E \approx 10^{24}$–10^{26}, dubbed *nanoflares*, can only be detected in EUV, because they seem not to exceed temperatures of $T \lesssim 2$ MK, and thus lack soft or hard X-ray emission. Frequency distributions reported in solar EUV have mostly concentrated on these nanoflares, but EUV statistics on larger flares are strangely lacking completely.

A first systematic study on EUV nanoflares was carried out by Krucker and Benz (1998), using images from the *Extreme-ultraviolet Imaging Telescope (EIT)* onboard the *SOlar and Heliospheric Observatory (SOHO)*. The detection of events in the EUV images was performed with a code similar to the one described in Section 6.9, using the 171 and 195 Å filters ($T \approx 1.1$–1.9 MK), and the energy was calculated based on a special physical model of the flare volume, assumed to be proportional to the area, i.e., $V \propto A$. This was the first study that reported significantly steeper powerlaw slopes ($\alpha_P \approx 2.3$–2.6) than

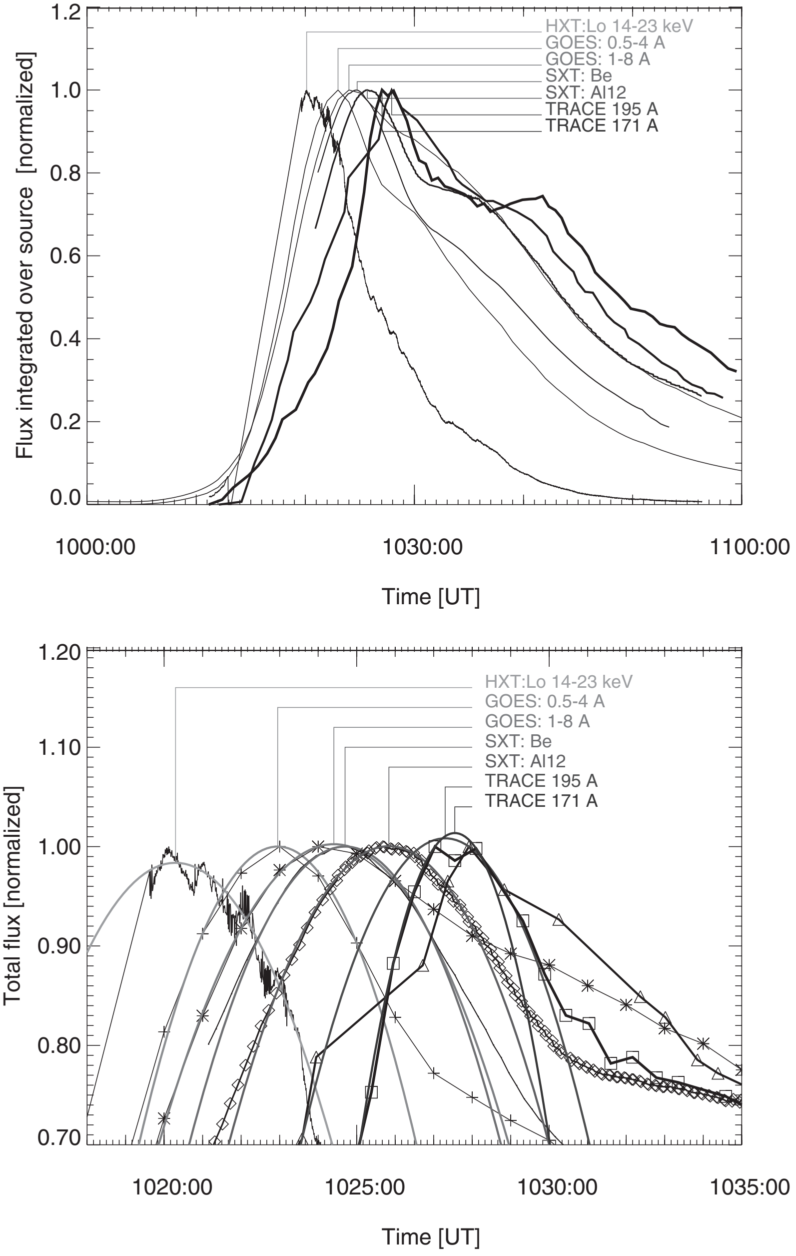

Fig. 7.12 *Top:* Light curves from Yohkoh/HXT (hard X-rays), Yohkoh/SXT and GOES (soft X-rays), and TRACE (EUV) of the 14 July 2000 Bastille-Day flare. Note that the different light curves are not self-similar. *Bottom:* Enlarged view of the emissions during their peak fluxes. Note a systematic delay that occurs in order of the decreasing temperature sensitivity of the instruments, due to the cooling of the flare plasma (Aschwanden and Alexander 2001).

Table 7.4 Frequency distributions measured in small-scale events in EUV, UV, and Hα. References: 1, Krucker and Benz (1998); 2, Aletti et al. (2000); 3, Parnell and Jupp (2000); 4, Aschwanden et al. (2000a,b); 5, Benz and Krucker (2002); 6, Aschwanden and Parnell (2002); 7, Georgoulis et al. (2002); 8, Greenhough et al. (2003); 9, McIntosh and Gurman (2005); 10, Nishizuka et al. (2009).

Powerlaw slope of peak flux α_P	Powerlaw slope of total fluence or energy α_E	Powerlaw slope of durations α_T	log range	Waveband λ (Å)	Reference
	$2.3-2.6$		1.3	171, 195	1
1.19 ± 1.13			2	195	2
	$2.0-2.6$		1.5	171, 195	3
$1.68-2.35$	1.79 ± 0.08		1.5	171, 195	4
	$2.31-2.59$		1.3	171, 195	5
	$2.04-2.52$		1.5	171, 195	5
1.71 ± 0.10	2.06 ± 0.10		2	171	6
1.75 ± 0.07	1.70 ± 0.17		2	195	6
1.52 ± 0.10	1.41 ± 0.09		1.5	AlMg	6
	1.54 ± 0.03		4	171+195+AlMg	6
2.12 ± 0.05			1	6563	7
$1.5-3.0$			1.5	1–500	8
		$1.4-2.0$	1	171,195,284	9
1.5		2.3	1.5	1550	10

previously reported in soft and hard X-rays. In Table 7.4 we list the powerlaw slopes α_E of energies (which are model-dependent), when the fluence was not reported. A similar study was done independently with another, but similar, event detection code, and powerlaw slopes in the range of $\alpha_P = 2.4$–2.6 were reported for the same volume model $V \propto A$, but a different range of $\alpha_P = 2.0$–2.1 for a modified volume model, i.e., $V \propto A^{3/2}$ (Parnell and Jupp 2000). The same data from these first two studies were reanalyzed with both volume models and powerlaw slopes of $\alpha_e = 2.52$–2.59 were inferred for the model $V \propto A$, and $\alpha_e = 2.04$–2.31 for the model $V \propto A^{3/2}$ (Benz and Krucker 2002), so we learned that the choice of the flare volume model changes the powerlaw slope of energies by about $\Delta\alpha \approx 0.4$.

A third study was conducted with the automated event detection code described in Section 6.9, which was designed to discriminate flare events (defined by impulsively heated and cooling loops) from non-flare events. For the frequency distributions of peak fluxes, a broken powerlaw was found at 171 Å with a slope varying from $\alpha_P = 1.68$ to $\alpha_P = 2.35$, but a single powerlaw at 195 Å with a slope of $\alpha_P = 1.85$ (Aschwanden et al. 2000a,b). Thermal flare energies were also determined using a cylindrical loop geometry for the flare volume, leading to a powerlaw slope of $\alpha_E = 1.79 \pm 0.08$ in the energy range of $E = 10^{24}$–$10^{26.5}$ erg (Fig. 1.14). This study also demonstrated that the selection of flare events can change the powerlaw slope by $\Delta\alpha \approx 0.3$. The next, more detailed, study was conducted using the combined EUV (TRACE) and soft X-ray data (Yohkoh), which allowed synthesize of a more complete temperature range, yielding more reliable total flare energies than previous studies in a single waveband (Aschwanden and Parnell 2002). The peak flux

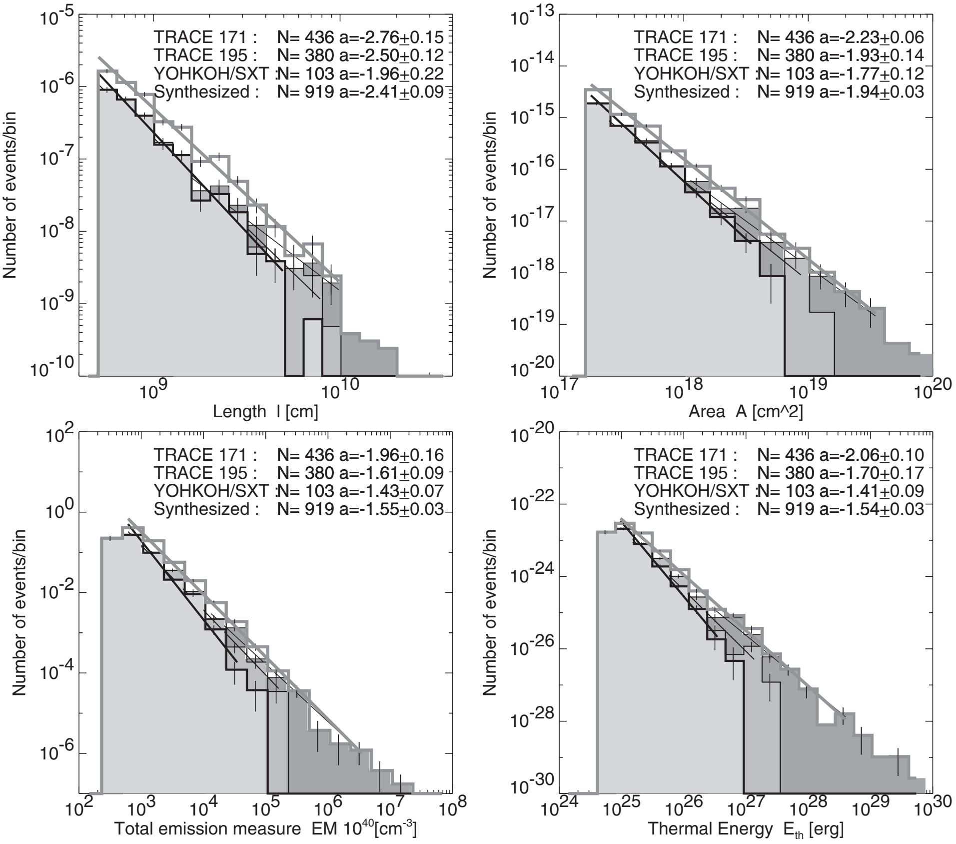

Fig. 7.13 Synthesized frequency distributions from all three wavebands (TRACE 171 Å, 195 Å, and Yohkoh/SXT AlMg) (gray histograms), along with the separate distributions from each waveband (in grayscales). Each of the distributions is fitted with a powerlaw, with the slope values and formal fit errors given in each panel. The four panels belong to the four parameters of the length l, area A, total emission measure M (which is proportional to the peak flux P), and the thermal energy E (Aschwanden and Parnell 2002).

distributions were found to have powerlaw slopes of $\alpha_P = 1.71 \pm 0.10$ (TRACE, 171 Å), $\alpha_P = 1.75 \pm 0.07$ (TRACE 195 Å), and $\alpha_P = 1.52 \pm 0.10$ (Yohkoh/SXT, AlMg, $T \gtrsim 2.4$ MK). Thermal flare energies were computed by taking the synthesized full temperature range as well as the fractal volume geometry (Chapter 8) of the flares into account, which yielded powerlaw slopes as steep as $\alpha_E = 2.06 \pm 0.10$ for the filter with the lowest temperature ($T \approx 1.0$ MK; TRACE 171 Å), or as low as $\alpha_E = 1.41 \pm 0.09$ for the filter with the highest temperature ($T \gtrsim 2.4$ MK; Yohkoh/SXT), while the synthesized distribution yields a powerlaw slope of $\alpha_E = 1.54 \pm 0.03$. Thus, this study demonstrated that there is also a temperature bias that steepens the powerlaw slope up to $\Delta\alpha \approx 0.9$, if the statistics is limited to a single narrowband filter of the lowest EUV temperature band. In Fig. 7.13 we show the frequency distributions of various parameters (length, area, total emission measure, and thermal energy) of the same event set measured in three different wavebands. Note the systematic flattening of the powerlaw slopes when including data with higher temperatures.

Further studies on frequency distributions were performed on EUV brightenings in the quiet Sun ($\alpha_P = 1.19 \pm 0.09$; Aletti et al. 2000), on full-disk EUV/XUV solar irradiance ($\alpha_P = 1.5$–3.0; Greenhough et al. 2003), or on EUV bright points over 9 years ($\alpha_T = 1.4$–2.0; McIntosh and Gurman 2005). Frequency distributions in flares were also evaluated for substructures that occur during a flare, such as UV brightenings of flare kernels observed in the C IV line ($\alpha_P = 1.5$, $\alpha_T = 2.3$; Nishizuka et al. 2009). Besides the EUV waveband, frequency distributions of small-scale variability events were also analyzed in the Hα line, which originates in the photosphere and chromosphere, such as in short-lived and small-scale events called *Ellerman bombs* ($\alpha_P = 2.12$; Georgoulis et al. 2002).

In summary we can say that the frequency distribution of nanoflares observed in EUV exhibit approximately the same powerlaw distributions of peak fluxes P and total energies E as observed in hard X-rays and soft X-rays, if the event definition is restricted to flare-like phenomena and if sufficiently broad temperature coverage is ensured to capture emission at the peak temperature of each event. However, several biases in the measurement of powerlaw slopes have been identified that are more severe in the cooler EUV waveband than in hotter soft X-ray wavebands, resulting from the event selection, narrowband temperature filters, and the geometric model of the flare volume, which enters the calculation of the thermal energy.

7.3.4 Solar Radio Emission

Solar radio bursts can be detected from ground-based instruments, and have thus been observed since their discovery by Hey and Southworth in 1942. Most solar radio bursts occur during solar flares, but they display a rich morphological variety that point to a number of different emission mechanisms, such as gyrosynchrotron emission of relativistic particles, electron beam-driven instabilities, loss-cone instabilities, or free-free (bremsstrahlung) emission. Radio emission at decimetric and microwave frequencies originate at the flare site and thus may show a detailed temporal co-evolution with the hard X-ray emission, while radio emission at metric and decametric wavelengths originate in the upper corona and heliosphere (Fig. 7.14), where they originate from local plasma instabilities or CME-driven shocks detached from the flare energy release process in the lower corona. This splits the statistics of solar radio bursts into two different realms, depending on the connectivity with the flare site, and consequently we expect possibly different frequency distributions for the two types. Solar radio bursts also span a large range of frequencies, from millimeter (≈ 300 GHz) to hectometer (≈ 3 MHz) wavelengths, and thus we might expect quite different frequency distributions depending on the wavelength or emission mechanism. A compilation of reported frequency distributions of solar radio bursts is given in Table 7.5.

The earliest frequency distributions of solar radio bursts were reported by Akabane (1956), who recorded solar radio bursts during 1951–1956 at 3 GHz and found powerlaw distributions with slopes of $\alpha_P \approx 1.8$. Further observations were reported in microwaves

Table 7.5 Frequency distributions measured from solar radio bursts, classified as type I storms, type III-like bursts, decimetric pulsation types (DCIM-P), decimetric millisecond spikes (DCIM-S), microwave bursts (MW), and microwave spikes (MW-S). References: 1, Akabane (1956); 2, Kundu (1965); 3, Kakinuma et al. (1969); 4, Fitzenreiter et al. (1976); 5, Aschwanden et al. (1995); 6, Aschwanden et al. (1998b); 7, Mercier and Trottet (1997); 8, Das et al. (1997); 9, Nita et al. (2002); 10, Ning et al. (2007).

Powerlaw slope of peak flux α_P	Powerlaw slope of total flux or total energy α_E	Powerlaw slope of durations α_T	log range	Waveband frequency f	Reference and type
1.8			2	3 GHz	1, MW
1.5			2	3, 10 GHz	2, MW
1.8			2	1, 2, 3.75, 9.4 GHz	3, MW
1.9–2.5			2	3.75, 9.4 GHz	3, MW
1.26–1.69			3	110 kHz–4.9 MHz	4, type III
1.28			2	100 MHz–3 GHz	5, type III
1.45 ± 0.31			3	100 MHz–3 GHz	6, type III
1.33 ± 0.11			3	100 MHz–3 GHz	6, DCIM-P
1.22–1.65			2.5	0.245–17 GHz	8, III, MW
1.71–1.91			4	0.100–2 GHz	9, III, MW
2.99 ± 0.63			3	100 MHz–3 GHz	6, DCIM-S
2.9–3.6			1.5	164, 237 MHz	7, type I
7.4 ± 0.4		5.4 ± 0.9	0.5	4.5–7.5 GHz	10, MW-S

(most likely to be produced by gyrosynchrotron emission) at 3 and 10 GHz during 1958–1959 with values of $\alpha_P = 1.5$ (Kundu 1965), at 1.2, 3.75, and 9.4 GHz during 1957–1962, with values of $\alpha = 1.8$ (Kakinuma et al. 1969), and at 3.75, and 9.4 GHz during 1957–1962, with values of $\alpha = 1.9$–2.5 (Kakinuma et al. 1969). Fitzenreiter et al. (1976) observed interplanetary type III bursts (produced by an electron beam instability) with the IMP-6 satellite during May–July 1971 at frequencies from 110 kHz to 4.9 MHz and found powerlaw distributions of their fluxes with slopes in the range of $\alpha_P = 1.26$–1.69. Interestingly, the value of the powerlaw slope systematically increases toward higher frequencies, which tells us something about the ratio of growth time τ_G to the saturation time t_S of the radio emission-producing instability, according to our model of exponentially growing instabilities, i.e. $\alpha_P = (1 + \tau_G/t_S)$ (Eq. 3.1.28). Statistics of flare-associated metric type III bursts yielded powerlaw slopes of $\alpha_P = 1.28$ (Aschwanden et al. 1995). Statistics on different types of flare-associated decimetric radio bursts included *decimetric type III types* (produced by electron beams) with $\alpha_P = 1.45 \pm 0.31$, *decimetric pulsation types* (produced by an oscillating instability) with $\alpha_P = 1.33 \pm 0.11$, and *decimetric millisecond spikes* (conceivably produced by an electron-cyclotron maser instability) with $\alpha_P = 2.99 \pm 0.63$ (Aschwanden et al. 1998b). There are always multiple radio bursts per flare, but when the distribution of peak fluxes or durations is investigated among the bursts occurring during a single flare, both powerlaw-like and exponential-like distributions are found. The relatively more restricted parameter space for a single flare could explain the exponential frequency distributions, which are not scale-free but define a dominant tem-

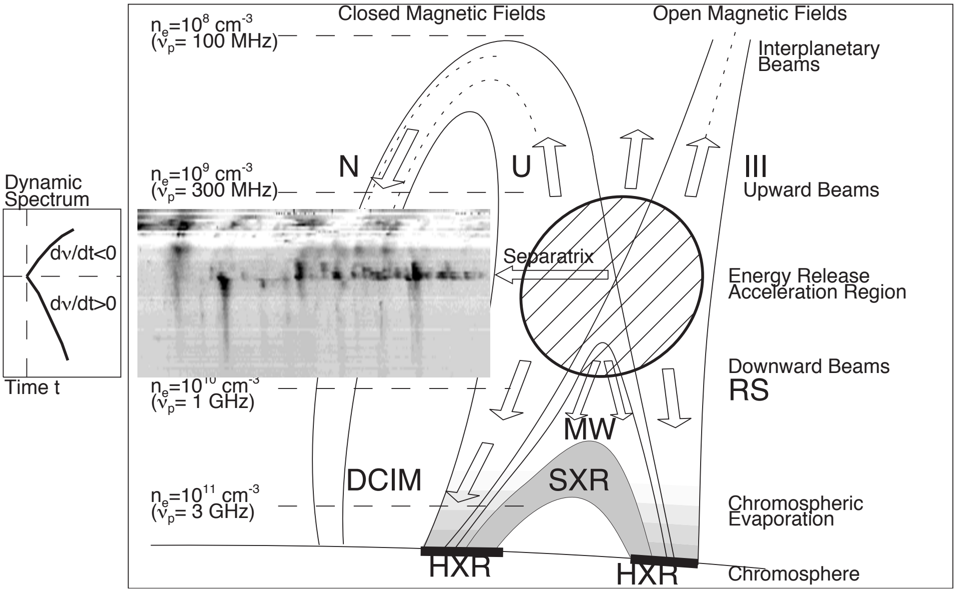

Fig. 7.14 Schematic overview of solar flare-related radio bursts: plasma emission excited by an electron beam instability produces radio bursts along open field lines escaping the acceleration region in upward direction (type III bursts), along upward escaping closed field lines (type U and N bursts), or in downward direction (reverse-slope drift [RS] bursts). Various decimetric (DCIM) radio bursts are produced by a losscone-type instability, sometimes with oscillatory patterns. Microwave emission (MW) produced by incoherent gyrosynchrotron emission mostly originates in flare loops where particles are injected from the acceleration region and subsequently become trapped (Aschwanden 2004).

poral or spatial scale. Mercier and Trottet (1997) sampled radio bursts from type I noise storms, which are produced above solar active regions without flares, probably associated with gentle continuous electron acceleration and found powerlaw distributions with slopes of $\alpha_P = 2.9$–3.6. Das et al. (1997) sampled radio bursts at frequencies from 245 MHz to 17 GHz and found some deviations from a strict powerlaw, which mostly affects the rollover at the low end of the distribution. A statistical analysis of decimetric millisecond spikes observed during single flares between 237 and 610 MHz exhibited both powerlaw-like and exponential-like flux distribution functions (Meszarosova et al. 1999, 2000). Isliker and Benz (2001) investigated how insufficient spatial and temporal resolution of these fast millisecond spikes affects the peak flux distribution function and found a tendency toward exponential behavior at large flux values. The most comprehensive statistics of solar radio bursts recorded over 40 years (1960–1999) compiled in NOAA catalogs was undertaken by Nita et al. (2002), finding powerlaw distributions with slopes in the range of $\alpha = 1.71$–1.91, using two different peak detection methods and sampling radio bursts in 8 frequency bands from 100 MHz to >2 GHz. Two examples of cumulative frequency distributions are shown in Fig. 7.15, measured at 2 GHz, which have a powerlaw slope (of the differential frequency distribution) of $\alpha = 1.82 \pm 0.01$ during the solar maximum and $\alpha = 1.81 \pm 0.02$ during the solar minimum, so there is no significant variation during the

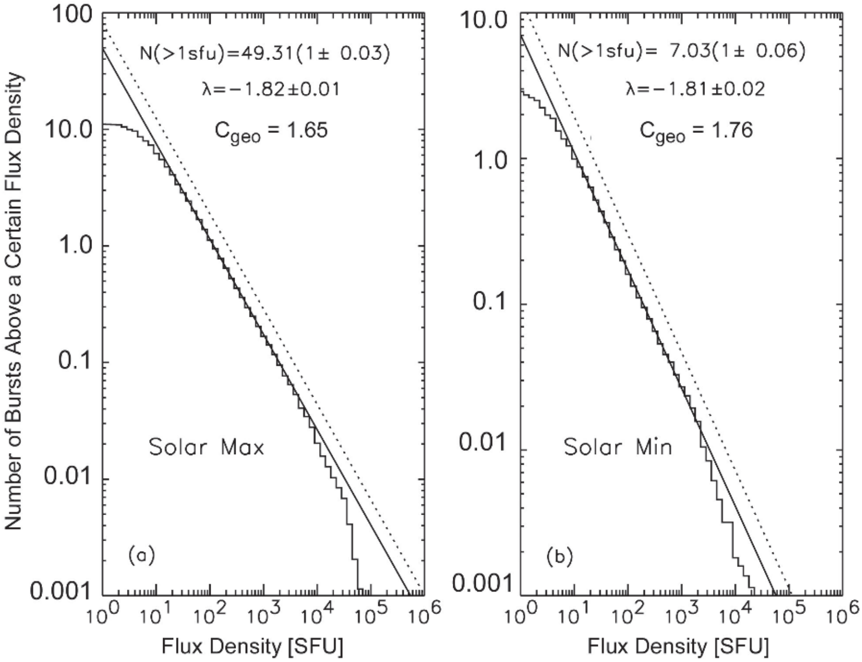

Fig. 7.15 Cumulative frequency distributions of radio bursts measured at 2 GHz during the solar maximum (left) and solar minimum (right) (Nita et al. 2002; reproduced by permission of the AAS).

solar cycle. Ning et al. (2007) reported very steep powerlaw slopes of $\alpha = 7.4 \pm 0.4$ for microwave bursts during a single event.

Inspecting Table 7.5 we see a clear pattern of two groups. The first group includes statistics of radio bursts occurring in many flare events, such as type III bursts and microwave bursts, which show similar powerlaw-like distributions as hard X-ray and soft X-ray flares, in the range of $\alpha_P \approx 1.3$–1.9. The second group includes statistics of radio burst fine structure during single flare events, such as decimetric millisecond spikes, type I sub-bursts, or microwave sub-bursts, which all exhibit very steep powerlaw distributions $\alpha \gtrsim 3$ or exponential distribution functions. The dissimilarity of statistical distributions of sub-bursts sampled during single flare events and the overall statistics sampled from many flares is clearly evident when they are juxtaposed in the same diagram (e..g, see examples in Aschwanden et al. 1998b). We conclude that physical parameters are more restricted during a single flare, and thus reveal a dominant temporal or spatial scale in the statistics of finestructure or sub-bursts, while a large statistical ensemble of many flares involves a much larger parameter range and produces the scale-free powerlaw distributions that are typical for flare statistics observed in other wavelength domains.

7.3.5 Solar Energetic Particle (SEP) Events

The highest particle energies detected in our heliosphere, mostly by in-situ detectors on spacecraft or by ground-based neutron monitors, can reach energies up to $\gtrsim$100 MeV for electrons and $\gtrsim$1 GeV for protons. While such high-energy particles were associated with cosmic rays earlier on, the current understanding is that they are accelerated either by shocks in coronal mass ejections, at typical distances of $R \approx 1$–5 solar radii, or in magnetic reconnection regions of solar flares in the lower corona. There is evidence for both scenarios, based on the timing inferred from the velocity dispersion of the detected particles: about half of the events have the time of their origin coincident with the flare peak times, while the other half originate with some significant delay, as expected for a CME-associated acceleration source. Nevertheless, whatever the origin of SEP events is, they represent very energetic phenomena and thus are expected to exhibit much flatter frequency distributions, like a subset of the largest and most energetic flare events.

An early frequency distribution of the intensity of 20–80 MeV protons (in units of protons cm^{-2} s^{-1} MeV^{-1}) was reported by Van Hollebeke et al. (1975), based on measurements of 185 SEP events during 1967–1972 with the *Interplanetary Monitoring Platform (IMP)* 4 and 5 spacecraft, who find a powerlaw distribution with a slope of $\alpha_P = 1.10 \pm 0.05$. Cliver et al. (1991) reported a powerlaw slope of $\alpha_P = 1.13 \pm 0.04$ for 24–43 MeV proton fluxes and $\alpha_P = 1.30 \pm 0.07$ for 3.6–18 MeV electron fluxes, based on 92 SEP events detected with the IMP-8 spacecraft during 1977-1983. Gabriel and Feynman (1996) collected data from the IMP 1, 2, 3, 5, 6, 7, 8, and the *Orbiting Geophysical Observatory (OGO)* 1 spacecraft observed during 1956–1990 and sampled frequency distributions of time-integrated particle fluxes (fluences), finding powerlaw slopes of $\alpha_E = 1.32 \pm 0.05$ for $>$10 MeV protons, $\alpha_E = 1.27 \pm 0.06$ for $>$30 MeV protons, and $\alpha_E = 1.32 \pm 0.07$ for $>$60 MeV protons, with little variation during the three solar cycles. A comprehensive compilation of the size distribution of $>$10 MeV solar proton events is provided by Miroshnichenko et al. (2001) for different datasets, based on the IMP spacecraft ($\alpha_P = 2.12 \pm 0.03$), the NOAA list ($\alpha_P = 1.47 \pm 0.06$), or SPE catalogues ($\alpha_P = 1.00$–1.43). The large range of powerlaw slopes in the latter dataset results from different threshold intensities, time ranges, or subsets with *sudden storm commencement (SSC)* associated events. The flattest slope $\alpha_P = 1.00 \pm 0.03$ was found for the lowest threshold, which corresponds to the most complete dataset, while the steepest powerlaw slope $\alpha_P = 2.12 \pm 0.03$ was found for the highest threshold and smallest data subset, and thus may be affected by the upper cutoff of the distribution. Looking at the temporal occurrence of SEP events, they are not well-correlated with the solar cycle, and thus unpredictable on time scales longer than the lifetime of an active region that has the necessary complex magnetic pattern (Xapsos et al. 2006; Hudson 2007).

In summary, the compilation in Table 7.6 shows that the frequency distributions of both the peak fluxes and fluences of SEP events are significantly flatter ($\alpha_P \approx \alpha_E \approx 1.1$–1.3) than for a comprehensive set of solar flares ($\alpha_P = 1.75 \pm 0.05$ and $\alpha_E = 1.61 \pm 0.04$). Since virtually all SEP events are also accompanied by a flare (unless the flare was occulted at the solar limb), the dataset of SEP events is essentially the most energetic subset of a flare distribution, regardless whether the acceleration of the high-energy particles occurred at

Table 7.6 Frequency distributions of solar energetic particle (SEP) events. References: 1, Van Hollebeke et al. (1975); 2, Belovsky and Ochelkov (1979); 3, Cliver et al. (1991); 4, Gabriel and Feynman (1996); 5, Smart and Shea (1997); 6, Mendoza et al. (1997); 7, Miroshnichenko et al. (2001); 8, Gerontidou et al. (2002).

Powerlaw slope of peak flux α_P	Powerlaw slope of total flux or total energy α_E	Powerlaw slope of durations α_T	log range	Energy range f	Reference and type
1.10 ± 0.05			3	20–80 MeV protons	1
1.40 ± 0.15				>10 MeV protons	2
1.13 ± 0.04			4	24–43 MeV protons	3
1.30 ± 0.07			4	3.6–18 MeV electrons	3
	1.32 ± 0.05		4	>10 MeV protons	4
	1.27 ± 0.06		4	>30 MeV protons	4
	1.32 ± 0.07		4	>60 MeV protons	4
1.47–2.42				>10 MeV protons	5
1.27–1.38				>10 MeV protons	6
1.00–2.12				>10 MeV protons	7
1.35			3.5	>10 MeV protons	8

the flare site in the lower corona or in associated CMEs further out in the heliosphere. The selection criterion for SEP events, e.g., the threshold for detecting > 10 MeV protons, includes all largest events of a peak flux distribution, but includes gradually less of the flare events with lower fluxes, which explains that their frequency distribution is flatter than for a complete set of flare events. Applying our exponential-growth model (Section 3.1), a powerlaw slope of $\alpha_P = (1+\tau_G/t_S) \approx 1.1$–$1.3$ corresponds to a mean ratio of $t_S/\tau_G = 3$–10 growth times, or mean amplification factors of $\exp(t_S/\tau_G) \approx 30$–$20{,}000$, which is hugely larger than the mean for average flares ($\exp(t_S/\tau_G) = \exp[1./(\alpha_E - 1)] \approx 5$. If we relate this mean amplification factor to the energy gain of the acceleration process, we expect that SEP events (with a powerlaw slope of $\alpha_E \approx 1.1$) produce a factor of $20{,}000/5 = 4{,}000$ higher energies, which explains the detection of > 10 MeV protons and >4 MeV electrons in SEP events.

7.4 Frequency Distributions in Astrophysics

While the Sun represents our local SOC laboratory that provides us abundant statistics and spatial information on each SOC event, observations of astrophysical sources offer only a few glimpses with sparse event statistics (due to the limited observing time allocation with expensive large telescopes) and no spatial information at all. This crucial limitation severely limits the characterization of frequency distributions, which requires ample statistics, but on the other hand, we can study exciting new phenomena that do not exist in our local solar system. The most-studied extra-solar SOC phenomena are stellar flares and accretion-disk or black-hole objects.

7.4.1 Stellar Flares

Here we review a few observations with published occurrence frequency distributions of stellar flares, mostly in UV wavelengths, which probe the transition regions at the base of stellar coronae. Robinson et al. (1999) observed the YZ Canis Minoris (YZ CMi) flare star (spectral type dM4.5e) on 1993 Nov 10 with the *High-Speed Photometer (HSP)* on the *Hubble Space Telescope (HST)* for 2.5 hrs and identified 54 flare events, finding a cumulative frequency distribution of the (time-integrated) flux with a slope of $\beta_E \approx 1.25 \pm 0.10$, approximately corresponding to a slope of $\alpha_E \approx 2.01 \pm 0.13$ for the differential frequency distribution, using Eq. (7.1.14) which includes the steepening effect of the cumulative frequency distribution near the upper cutoff, based on the logarithmic range $E_{max}/E_{min} \approx 10$ (see Fig. 3 in Robinson et al. 1999). The cumulative frequency distribution exhibits two bumps, so it is not well-characterized by a powerlaw function, which is often the case for small samples.

Audard et al. (2000) sample the flare activity of 12 (late-type) cool stars (spectral type F to M) from *Extreme Ultraviolet Explorer (EUVE)* Deep Survey observations. The cumulative frequency distributions of their total radiative energies E (which is assumed to be proportional to the total number of photon counts observed in the energy range of 0.01–10 keV) of these 12 stars is shown in Fig. 7.16. The cumulative frequency distributions are shown in form of rank-order plots, since there are only about 5–15 datapoints (flare events) measured for each star. Fitting a powerlaw distribution in the log-log plane they find the cumulative powerlaw indices β and estimate the powerlaw indices of the differential frequency distribution with the approximate relation $\alpha = \beta + 1$, listed in the first column of Table 7.7 (which corresponds to the value α^b in the fourth column of their Table 2). We estimate the corresponding powerlaw slopes α of the differential frequency distributions with the relationship $\alpha(\beta)$ given in Eq. (7.1.14), based on a powerlaw fit in the lower half of the cumulative distribution (marked with a thick line in Fig. 7.16) and the logarithmic range $\Delta E_{log} = \log(E_{max}/E_{min})$, with $q_2 = 10^{(-\Delta E_{\log}/2)}$, listed as values α_E^P in Table 7.7. We also fit the exact cumulative distribution function $N^{cum}(>x)$ as defined in Eq. (7.1.10), which includes the steepening at the upper end (marked with thin curves in Fig. 7.16), listed as values α_E^C in Table 7.7. We see that discrepancies between the three methods mostly arise where the logarithmic range of the sampled energies is small (listed in parenthesis in the column ΔE_{log} in Table 7.7), say $\lesssim 0.7$ decades (a factor of 5), excluding the rightmost energy bin containing the largest event. Powerlaw fits over such small ranges are not reliable because they could fit the gradual exponential-like cutoff without constraining the powerlaw part at lower energies. Thus, if we focus on the more reliable events with energy ranges of $\Delta E_{log} \geq 0.8$ (excluding the uppermost bin), we find 7 cases (out of the 12 stars analyzed by Audard et al. 2000) which have the following mean powerlaw slopes for each of the three methods: $\beta + 1 = 2.01 \pm 0.15$, $\alpha_E^P = 1.65 \pm 0.18$, and $\alpha_E^C = 1.75 \pm 0.26$. Thus, we find a significantly flatter slope of $\alpha_E \approx 1.7 \pm 0.2$ based on the two methods (α_E^P and α_E^C) that include the upper cutoff effect in the cumulative frequency distribution than the method $\alpha = \beta + 1$ that neglects this effect, used in Audard et al. (2000). A summary of various biases in the derivation of frequency distributions from stellar data is given in Güdel et al. (2003).

Table 7.7 Frequency distributions observed in stellar flares. The powerlaw slope α of the differential frequency distribution of energies is calculated with three methods: $\beta+1$ is from a powerlaw fit to the cumulative distribution (reported by authors); α_E^P is from a powerlaw fit to the lower half of the bins with correction Eq. (7.1.14), and α_E^C by fitting the cumulative distribution (Eq. 7.1.10). The logarithmic ranges $\Delta E_{log} = \log(E_{max}/E_{min})$ are given, where the numbers in parentheses give the range excluding the uppermost bin in the cumulative distribution that contains the largest flare. The values flagged with an asterisk were obtained from fitting photon arrival time distributions using Monte-Carlo simulations. References: 1, Robinson et al. (1999); 2, Audard et al. (2000); 3, Kashyap et al. (2002); 4, Güdel et al. (2003); 5, Arzner and Güdel (2004); 6, Arzner et al. (2007); 7, Stelzer et al. (2007).

Powerlaw slope of total flux $\beta+1$	α_E^P	α_E^C	Logarithmic range ΔE_{log}	Object	Instrument	Reference
2.25 ± 0.10	2.01		1.0	YZ Cmi	HSP/HST	1
1.89	2.40	2.15	1.6 (0.6)	HD 2726	EUVE	2
1.98	1.79	1.95	1.3 (0.8)	47 Cas	EUVE	2
2.27	1.71	1.93	1.0 (0.9)	EK Dra	EUVE	2
1.90	1.67	2.14	0.8 (0.4)	κ Cet 1994	EUVE	2
2.21	2.22	2.18	1.0 (0.6)	κ Cet 1995	EUVE	2
1.97	1.64	1.83	1.4 (1.1)	AB Dor	EUVE	2
2.50	2.00	2.68	0.5 (0.4)	ε Eri	EUVE	2
1.96	1.65	1.69	1.3 (1.1)	GJ 411	EUVE	2
1.85	1.72	1.81	1.5 (1.2)	AD Leo	EUVE	2
1.90	1.75	1.84	1.3 (1.0)	EV Lac	EUVE	2
1.91	2.20	1.79	1.6 (0.7)	CN Leo 1994	EUVE	2
2.14	1.26	1.19	1.0 (0.8)	CN Leo 1995	EUVE	2
2.60 ± 0.34*				FK Aqr	EUVE	3
2.74 ± 0.35*				V1054 Oph	EUVE	3
2.03–2.32*				AD Leo	EUVE	3
2.0–2.5*				AD Leo	EUVE	4
2.3 ± 0.1*				AD Leo	EUVE	5
1.9–2.5*				HD 31305	XMM	6
2.4 ± 0.5				TMC	XMM	7

Kashyap et al. (2002) analyzed also observations from the EUVI Deep Survey and inferred the frequency distribution $N(E) \propto E^{-\alpha}$ in an indirect way by Monte–Carlo simulations of photon arrival times, where the numerical model has three free parameters: α the powerlaw index of the energy distribution, r_F the average count rate due to flares, and r_C the average background count rate. So, the value α is found from the best fit of the modeled to the observed distribution of photon arrival times. The obtained values in the range of $\alpha \approx 2.2$–2.7 are significantly steeper than previously inferred values from similar stars. A similar value was obtained for AD Leo using the same method (Güdel et al. 2003; Arzner and Güdel 2004). Arzner et al. (2007) applied the same Monte-Carlo simulation technique to a sample of 22 stars observed with the *XMM-Newton* Extended Survey of the Taurus Molecular Cloud (XEST), but could constrain the powerlaw slope of the flare energy distribution with an acceptable fit ($\alpha_E = 2.0_{1.9}^{2.5}$) only for one case (HD 31305). It would be interesting to test the validity of this novel simulation method (that infers powerlaw slopes from fitting photon arrival time distributions) by comparing with powerlaw

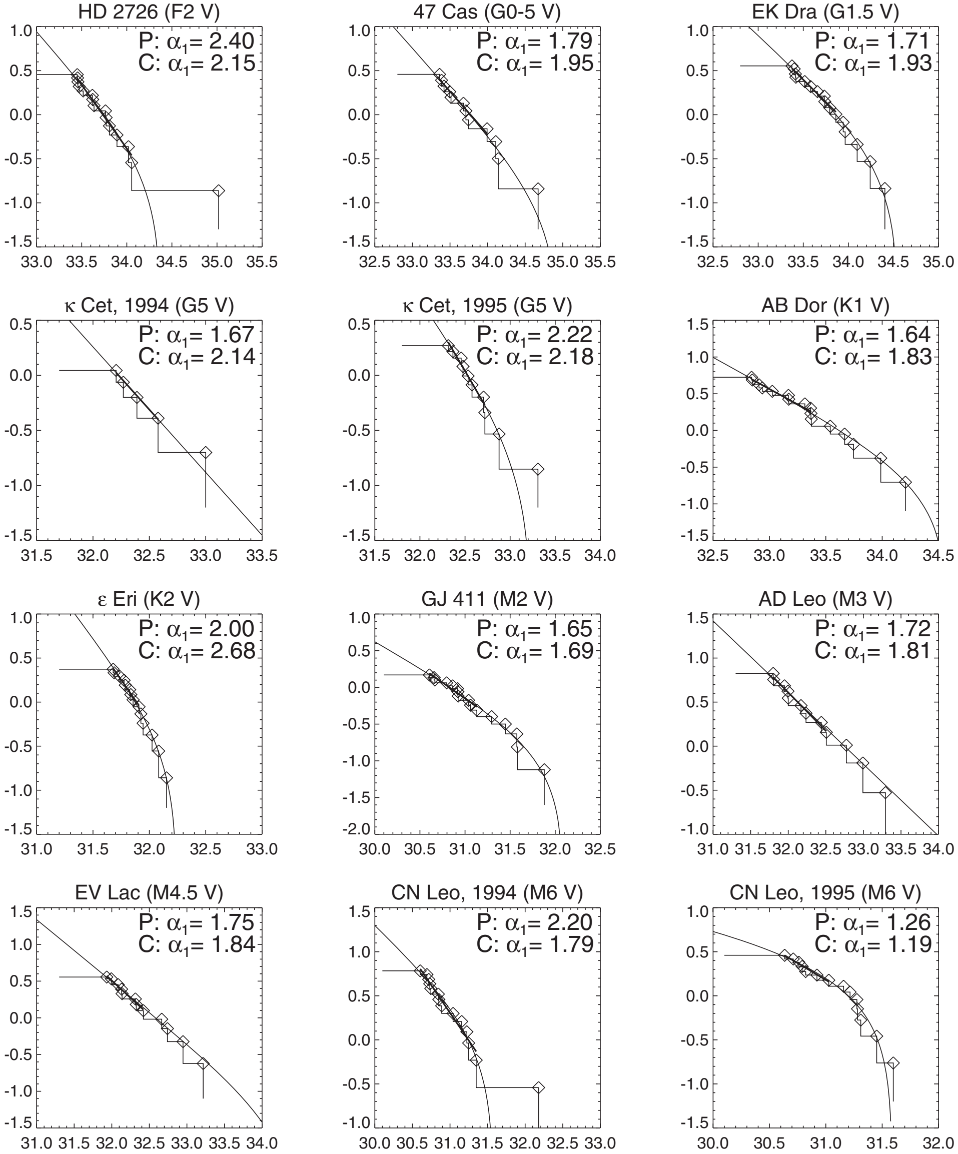

Fig. 7.16 Cumulative frequency distributions of flare energies (total counts) observed for 12 cool (type F to M) stars with EUVE (Audard et al. 2000). The flare events are marked with diamonds, fitted with a powerlaw fit in the lower half (P; thick line), and fitted with a cumulative frequency distribution (C; curved function).

slopes obtained from cumulative frequency distributions using the same data sets. Systematic biases of this method are not known yet, but it is conceivable that some assumptions (e.g., self-similar flare time profile) could influence the inferred powerlaw slopes (Arzner et al. 2007).

The inference of powerlaw slopes of frequency distributions from stellar flares cannot be obtained in the same way as for solar flares, where abundant statistics is available and a powerlaw slope can directly be fitted to the differential frequency distribution. Instead, the very small samples of flaring events per star require either the inversion of a rank-order plot (or cumulative frequency distribution) or a Monte Carlo simulation technique that fits a distribution of observed photon arrival times. Both methods have their own bias that need to be determined. For cumulative frequency distributions, the effect of the upper cutoff needs to be taken into account in small samples, which changes the powerlaw slope in the order of $\Delta\alpha \approx 0.3$. If this effect is taken into account, we infer values of $\alpha_E \approx 1.7 \pm 0.2$, which is similar to solar flares $\alpha_E \approx 1.61 \pm 0.04$ (Fig. 7.8), although the energies of the detected stellar flares are up to two orders of magnitude higher than the largest solar flares (Fig. 1.15).

7.4.2 Pulsar Glitches

Pulsars exhibit glitches in pulse amplitudes and frequency shifts that correspond to large positive spin-ups of the neutron star, probably caused by sporadic unpinning of vortices that transfer momentum to the crust. Conservation of the momentum produces then an increase of the angular rotation rate, like a twirling ice skater who draws the hands closer. The pulse height distribution of the Crab pulsar (NGC 0532 or PSR B0531+21) observed at 146 MHz was found to have a powerlaw slope of $\alpha_P \approx \beta + 1 = 3.5$ over a range of 2.25 to 300 times the average pulse size (Argyle and Gower 1972). Similar values were measured by Lundgren et al. (1995), with $\alpha_P \approx 3.06$–3.36 (Fig. 7.17). While the Crab pulsar is the youngest known pulsar (born in the year 1054), PSR B1937+21 is an older pulsar with a 20 times faster period (1.56 ms) than the Crab pulsar (33 ms). Cognard et al. (1996) measured a powerlaw distribution with a slope of $\alpha_P \approx \beta + 1 = 2.8 \pm 0.1$ from its occasional giant pulses. A theoretical model of the inertial momentum change in pulsar macro-glitches predicts a frequency distribution of $N(E) \propto E^{-1.14}$ (Morley and Garcia-Pelayo 1993), which is much flatter than previously observed. However, statistics on nine pulsars found powerlaw slopes in a range of $\alpha_E = -0.13, \ldots, 2.4$ (see Table 7.8) for the size distribution of pulse glitches (Melatos et al. 2008), but no correlation between the powerlaw slope and the spin-down age was found.

Interestingly, while most pulsars have a Gaussian or exponential pulse-amplitude distribution, only few pulsars, including the Crab pulsar, exhibit a powerlaw (Lundgren et al. 1995), which could be interpreted in terms of a SOC phenomenon (Young and Kenny 1996). Assuming a SOC model would also imply a powerlaw distribution for the pulse duration. Turbulence in neutral non-ionized fluids were considered as a possible mechanism that exhibit spatial and temporal scale invariance (Young and Kenny 1996). Alternative models in terms of modulational instabilities with stochastic growth and wave collapse were also proposed, which produce log-normal energy distributions with a steep power-

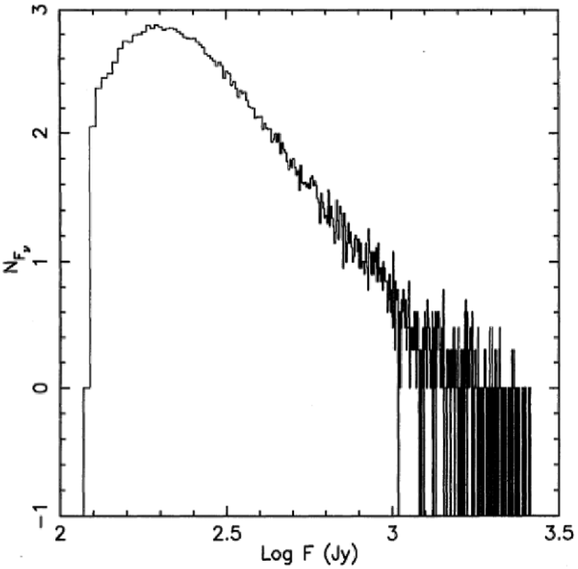

Fig. 7.17 Frequency distribution of giant-pulse flux densities measured from the Crab pulsar, observed during 15–27 May 1991 with the Green Bank 43-m telescope at 1,330, 800, and 812.5 MHz. The tail can be represented by a powerlaw distribution $N_F \propto F^{-\alpha}$ with $\alpha = 3.46 \pm 0.04$ for fluxes $F > 200$ Jy (Lundgren et al. 1995; reproduced by permission of the AAS).

Table 7.8 Frequency distributions observed from pulsar (giant-pulse) glitches (PSR), soft gamma-ray repeaters (SGR), black-hole object Cygnus X-1, and blazar GC 0109+224. Uncertainties in terms of one standard deviation are quoted in parentheses [...] for some cases. References: 1, Argyle and Gower (1972); 2, Lundgren et al. (1995); 3, Cognard et al. (1996); 4, Melatos et al. (2008); 5, Gogus et al. (1999); 6, Gogus et al. (2000); 7, Chang et al. (1996); 8, Mineshige and Negoro (1999); 9, Ciprini et al. (2003).

Powerlaw slope flux α_P	Powerlaw slope fluence α_E	Waveband	Object	Ref.
3.5		146 MHz	Crab pulsar	1
3.06-3.36		813–1330 MHz	Crab pulsar	2
2.8 ± 0.1		430 MHz	PSR B1937+21	3
2.4 [1.5,5.2]			PSR 0358+5413	4
1.2 [1.1,1.4]			PSR 0534+2200	4
0.42 [0.39,0.43]			PSR 0537-6910	4
1.8 [1.2,2.7]			PSR 0631+1036	4
−0.13 [−0.20,+0.18]			PSR 0835-4510	4
1.4 [1.2,+2.1]			PSR 1341-6220	4
1.1 [0.98,1.3]			PSR 1740-3015	4
0.57 [0.092,1.1]			PSR 1801-2304	4
0.36 [-0.30,1.0]			PSR 1825-0935	4
	1.66	>25 keV	SGR 1900+14	5
	1.43, 1.76, 1.67	20.8 keV	SGR 1806-20	6,7
7.1		1.2–58.4 keV	Cygnus X-1	8
1.55		optical	GC 0109+224	9

law tail $\alpha_E \approx 4$–7 at high energies, which could correspond to the "giant pulses" (Cairns 2004; Cairns et al. 2004).

Finally, a detailed cellular automaton SOC model was proposed for pulsar glitches, which could reproduce powerlaw slopes of $\alpha_E \approx 2.0$–4.3 for pulse sizes and $\alpha_T = 2.2$–5.5 for pulse durations (Warzawski and Melatos 2008). The underlying theoretical model of pulsar glitches is summarized by Warzawski and Melatos (2008) as follows: *The neutron superfluid in the stellar interior is threaded by many ($\approx 10^{16}$) vortices, approximately one per cent of which are pinned to the stellar crust at grain boundaries and/or nuclear lattice sites. As the pulsar crust spins down electromagnetically, a lag builds up between the velocity of the pinned vortex lines (corotating with the crust) and the superfluid. When the transverse Magnus force (directly proportional to the lag) surpasses a threshold value (equal to the strength of the pinning force), a catastrophic unpinning of vortices occurs, transferring angular momentum to the crust. In order for this mechanism to generate glitches on the scale observed, it requires up to 10^{12} vortices to unpin simultaneously, exhibiting a high level of collective, non-local behaviour.*

7.4.3 Soft Gamma-Ray Repeaters

Observations with the *Compton Gamma Ray Observatory (CGRO)* revealed a rare class of objects that show repetitive emission of low-energy gamma rays (>25 keV), termed *soft gamma-ray repeaters (SGR)*. In 1999, only four such SGR sources were known (three in our galaxy and one in the Magellanic Cloud), but at least three of them were identified to be associated with slowly rotating, extremely magnetized neutron stars, located in supernova remnants (Kouveliotou et al. 1998, 1999). They emit gamma-ray bursts with relatively soft spectra (like optically-thin bremsstrahlung at $k_B T \approx 20$–40 keV) and short duration of $\approx$0.1 s. Thompson and Duncan (1996) suggested that these gamma-ray bursts occur from neutron star crust fractures driven by the stress of an evolving, ultrastrong magnetic field ($B \gtrsim 10^{14}$ G).

Gogus et al. (1999) analyzed a database of 187 gamma-ray bursts (at energies of $\geq$25 keV) from SGR 1900+14 during the 1998–1999 active phase and found that the fluence or energy distribution of the bursts follows a powerlaw distribution over 4 orders of magnitude (Fig. 7.18, left). Also a correlation between the energy and duration was found, $E \propto T^{1.13}$, similar to solar flares (Eq. 7.3.2). Gogus et al. (2000) analyzed 290 events of SGR 1806-20 using data from the *Rossi X-Ray Timing Explorer (RXTE)*, 111 events detected with *CGRO/BATSE*, and 134 events detected with the *International Cometary Cometary Explorer (ICE)*, and found powerlaw slopes of $\alpha_E = 1.43$, 1.76, and 1.67 for the fluences, respectively (Fig. 7.18, right). The results were interpreted in support of the neutron star crustquake model of Thompson and Duncan (1996), in analogy to the SOC interpretation of earthquakes.

Soft gamma-ray repeaters with pulses originating from the same object are the exception rather than the rule, while most gamma-ray bursts detected with CGRO are non-repetitive, and thus come sporadically from different objects. Statistics of the temporal properties of those gamma-ray bursts has been gathered for several hundreds of events (e.g., Norris 1995; Norris et al. 1996; Quilligan et al. 2002), but it is not clear whether any

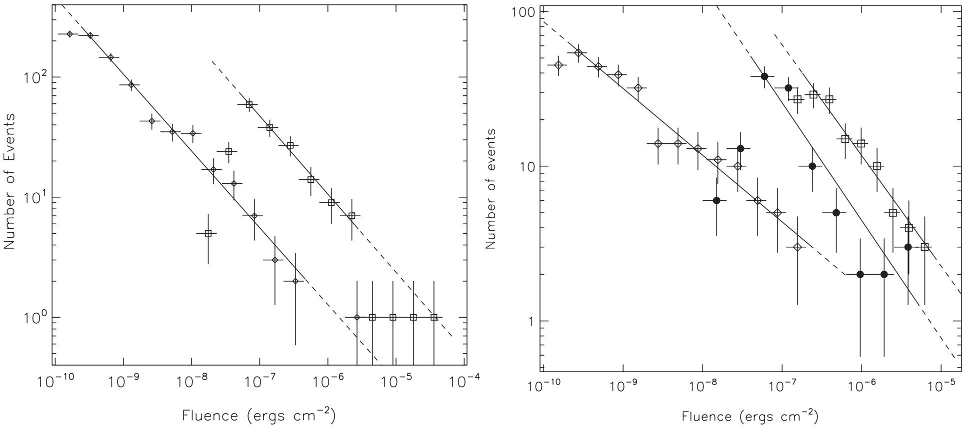

Fig. 7.18 Differential frequency distributions of the fluences of soft gamma-ray repeater sources: SGR 1900+14 (left), and SGR 1806-20 (right), observed with CGRO, RXTE, and ICE (Gogus et al. 1999, 2000; reproduced by permission of the AAS).

SOC characteristic could be retrieved from an observational sample that contains only one event per SOC system, where each SOC system has vastly different (unknown) distances to the observer.

7.4.4 Black Hole Objects

We discussed observations of black-hole candidates such as Cygnus X-1 in Section 1.9, numerical cellular automaton models of the surrounding accretion disks in Section 2.7.1, a shot-noise model of their power spectra in Section 4.8.4, and their waiting-time distributions in Section 5.7.2. In Fig. 7.19 (right) we show an observed occurrence frequency distribution of the peak intensity of the shots from a light curve of Cygnus X-1 (Negoro et al. 1995; Mineshige and Negoro 1999), along with a theoretical distribution (Fig. 7.19, left), simulated according to the cellular automaton model of Mineshige et al. (1994a,b). The observed peak-intensity distribution has a steep slope of approximately $\alpha_P \approx 7.1$. The cellular automaton model can accommodate a range of powerlaw slopes, depending on what fraction of mass m' (Eq. 2.7.4) is transferred by gradual diffusion in addition to the avalanche-like shots, e.g., simulations with $m' = m/100, m/10$, or $m/5$ produce powerlaw slopes of $\alpha_P \approx 5.6, 7.7$, and 11.5 (Mineshige and Negoro 1999; Takeuchi et al. 1995). Initial simulations without gradual diffusion produced energy distributions with powerlaws of $N(E) \propto E^{-2.8}$ and time scale distributions of $N(T) \propto T^{-1.4}$ (Mineshige et al. 1994b). Whatever the detailed scaling of the mass transfer in the accretion disk is, the fact of a powerlaw distribution of peak fluxes in the light curve is thought to support a SOC interpretation in terms of an avalanching system in a self-organized critical state.

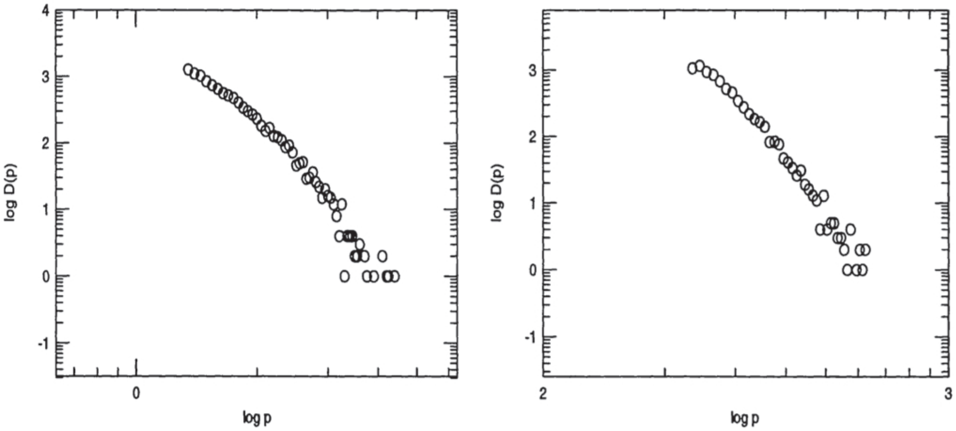

Fig. 7.19 *Left:* Numerically simulated frequency distribution of a cellular automaton model of mass avalanches in an accretion disk (Mineshige and Negoro 1999). *Right:* Observed frequency distribution of the peak intensities of pulses in the light curve of the black-hole object Cygnus X-1, exhibiting a powerlaw slope of $\alpha_P \approx 7.1$ (Negoro et al. 1995; Mineshige and Negoro 1999).

7.4.5 Blazars

Blazars (blazing quasi-stellar objects) are very compact quasars (quasi-stellar objects) associated with super-massive black holes in the center of active, giant elliptical galaxies. They represent a sub-group of active galactic nuclei (AGNs) which emit a relativistic jet in the direction of the Earth. Because of this particular geometry, where the jets are co-aligned with the line-of-sight to the observer, rapid variability and apparent super-luminous features are the paramount characteristics of these objects.

The optical variability of blazar GC 0109+224 was monitored from 1994 and the light curve was found to exhibit an intermediate behavior between flickering and shot noise, with a power spectrum of $P(\nu) \propto \nu^{-p}$ with $1.57 < p < 2.05$ (Ciprini et al. 2003). A combination of two modes between flickering (pink noise with $p > 0.8$) and pure shot noise (Brownian random walk or brown noise with $p \geq 2$) seems to be common in blazars (Hufnagel and Bregman 1992). Ciprini et al. (2003) constructed an occurrence frequency distribution of the peak fluxes of flare events and found a powerlaw distribution $N(P) \propto P^{-\alpha}$ with a slope of $\alpha \approx 1.55$, within a range of about one order of magnitude, and excluding the largest flares (Fig. 7.20). The powerlaw distribution of peak fluxes, along with the $1/f$ flicker noise spectrum of the light curve, was considered as an indication that blazars also represent a SOC phenomenon (Ciprini et al. 2003).

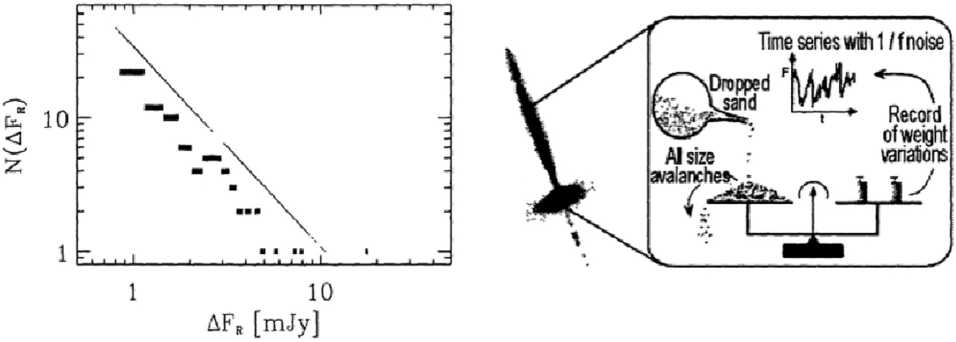

Fig. 7.20 *Left:* Frequency distribution of peak fluxes of flaring events in blazar GC 0109+224, including fluxes above of a 3σ-threshold, fitted with a powerlaw $N(P) \propto P^{-1.55}$. *Right:* A cartoon that illustrates the analogy of toppling avalanches in SOC sandpiles with jets emerging out of a blazar (Ciprini et al. 2003).

7.5 Summary

We have reviewed most of the occurrence frequency distributions observed in astrophysical event sets: magnetospheric substorms, solar and stellar flares, solar energetic particle events, solar radio bursts, pulsar glitches, soft gamma-ray repeaters, black hole objects, and blazars. Many frequency distributions of peak fluxes or fluences are found to be close to powerlaw distributions, with slopes varying in a considerable range of $\alpha \approx 1,...,10$, with a preference around $\alpha \approx 1.5$–2.0 for most phenomena. Statistics of the same phenomenon type exhibit their own characteristic value, such as (in increasing order): $\alpha_E \approx 1.1$–1.3 for magnetospheric substorm events, $\alpha_P \approx 1.1$–1.5 for solar energetic particle events (SEP), $\alpha_E \approx 1.4$–1.8 for soft gamma-ray repeaters, $\alpha_P \lesssim 1.5$ for blazars, $\alpha_P \approx 1.5$–1.8 for solar radio bursts, $\alpha_E \approx 1.6$–1.8 for solar (and probably stellar) flares, $\alpha_E \approx 3$ for pulsar glitches, or $\alpha_P \approx 7$ for black-hole objects. We identified a number of measurement biases that entered the published values, such as: (1) the "instrumental waveband bias" and "incomplete temperature coverage bias", which can lead to an overestimate of the powerlaw slope (e.g., for solar nanoflares detected with narrowband EUV filters); (2) the "big-event selection bias", which can lead to an underestimate of the powerlaw slope for event subsets that select larger events with a higher probability (e.g., SEP or CME events are not representative subsets of solar flares); or (3) the "upper-cutoff bias of cumulative frequency distributions", which leads to an overestimate of the powerlaw slope for small samples (e.g., stellar flares). All these biases in the measurement of powerlaw slopes and derivation of power indices in parameter correlations can be systematically studied with Monte-Carlo simulations (Section 7.1.4) and forward-fitted to the observed data. A self-consistent determination of the powerlaw slopes of peak fluxes (α_P), fluences (α_E), and durations (α_T) can quantify the correlations between the observables (P, E, T) that are most important for infering the scaling laws of underlying physical processes (see Chapter 9).

7.6 Problems

Problem 7.1: Simulate a distribution of random values that obey an exponential frequency distribution and verify that the histogrammed differential frequency distribution matches the analytical exponential function (follow Eqs. 7.1.28–7.1.30 and Fig. 7.3, left).

Problem 7.2: Simulate a distribution of random values that obey a powerlaw frequency distribution and verify that the histogrammed differential frequency distribution matches the analytical powerlaw function (follow Eqs. 7.1.31–7.1.33 and Fig. 7.3, right).

Problem 7.3: Use the numerically generated values of Problem 7.2 to construct the cumulative frequency distribution function and a rank-order plot (Fig. 7.2). What powerlaw slope do you infer from the cumulative frequency distribution or rank-order plot and how do you explain the difference to the original powerlaw slope of the numerically generated values?

Problem 7.4: Simulate the third case ($n = 10^4$ events) shown in Fig. 7.4 with different sets of random numbers and quantify the average accuracy or reproducibility or the powerlaw slopes and power indices of the parameter correlations.

Problem 7.5: What are necessary and sufficient conditions that the frequency distributions of solar flare energies observed in soft X-ray and hard X-ray wavelengths be identical?

Problem 7.6: Simulate a small sample of 15 random events that obey a powerlaw distribution function with a slope of $\alpha_E = 1.5$ to mimic a dataset of stellar flares (Fig. 7.16) and determine the powerlaw slope with three different methods: (1) with an overall powerlaw fit, (2) with a half powerlaw fit and the correction given in Eq. (7.1.14), and (3) with fitting the cumulative distribution function given in Eq. (7.1.10). How much different are the values determined with the three methods? Do you find a systematic bias when repeating the same experiment with different random number sets?