5.3.4 Maximum likelihood estimation

The procedure described in Sec. 4.4.5 provides a recipe for finding the most likely (least biased) set of probabilities, by maximizing the "entropy" functional

$$S[\{p_i(E_i)\}] = -\sum_{i=1}^{\ell} \int d\Gamma_i p_i(E_i) \ln p_i(E_i) \,, \tag{5.3.15}$$

where $d\Gamma_i$ indicates integration over all internal coordinates contributing to energy, e.g. x and v for one particle with energy $E = V(x) + mv^2/2$.[7] The maximization has to be done subject to the constraint on energy from Eq. (5.3.14):

$$\sum_{i=1}^{\ell} \int d\Gamma_i p_i(E_i) E_i = \text{constant} \,, \tag{5.3.16}$$

as well as the normalization requirements

$$\int d\Gamma_i p_i(E_i) = 1 \quad \text{for all } i \,. \tag{5.3.17}$$

To enforce the constraints we introduce a *Lagrange multiplier* β for Eq. (5.3.16), and the set $\{\alpha_i\}$ for Eqs. (5.3.17). Including these constraints, we need to maximize the functional

$$S[\{p_i(E_i)\}] = -\sum_{i=1}^{\ell} \int d\Gamma_i p_i(E_i) \left[\ln p_i(E_i) + \beta E_i + \alpha_i\right] \,. \tag{5.3.18}$$

Setting the functional derivative to zero gives

$$\frac{\delta S}{\delta p_i(E_i)} = -\left[\ln p_i(E_i) + \beta E_i + \alpha_i\right] - 1 = 0 \,. \tag{5.3.19}$$

The solution thus has the form

$$p_i(E_i) = \frac{1}{Z_i} e^{-\beta E_i} \,, \tag{5.3.20}$$

where the parameters $Z_i = e^{\alpha_i + 1}$ have to adjusted for proper normalization. The important part is that the same parameter β governs the exponential distribution of energies of all bodies in the bath.

The parameter β is an intrinsic property of the bath which governs the equilibrium amongst bodies with which it exchanges energy. A "hot" bath with $\beta \to 0$ imposes almost no constraint on the energy of each body, while a "cold" bath as $\beta \to \infty$ restricts the energies E_i to smallest possible values. In statistical mechanics this property of the bath is related to its absolute temperature T through $\beta = 1/(k_B T)$, where k_B is the Boltzmann constant, and the exponential form of the PDF is referred to as the *Boltzmann weight*.

[7] Since changing coordinates modifies the expression for entropy, there is some ambiguity involved here. It can be shown that the correct choice of coordinates corresponds to pairs of canonically conjugate variables such as x and $p = mv$.