In his 1917 paper, "On the Quantum Theory of Radiation," Einstein uses the following argument to derive Planck's law. Readers should note that the idea of induced emission was a theoretical discovery by Einstein. Prior to Einstein, no one was aware that induced emission existed.

Consider a gas at temperature $T$. Let $N$ be the number of molecules in the gas and let $N_n$ be the number of molecules with internal energy $\varepsilon_n$. By the Maxwell-Boltzmann distribution we have

$$\frac{N_n}{N} = p_n \exp\left(-\frac{\varepsilon_n}{kT}\right)$$

where $k$ is Boltzmann's constant. The coefficient $p_n$ is a statistical weighting factor that does not depend on $T$.

Let us now consider the processes by which a molecule transitions between energy levels. The processes are absorption, induced emission, and spontaneous emission. Let $\varepsilon_m$ be an energy level such that $\varepsilon_m > \varepsilon_n$. Let $B_n$, $B_m$, and $A_m$ be coefficients of transition rates such that

$$\underset{\substack{\text{absorption}\\ \varepsilon_n \to \varepsilon_m}}{\frac{dN_n}{dt} = B_n \rho(\nu) N_n} \qquad \underset{\substack{\text{induced emission}\\ \varepsilon_m \to \varepsilon_n}}{\frac{dN_m}{dt} = B_m \rho(\nu) N_m} \qquad \underset{\substack{\text{spontaneous emission}\\ \varepsilon_m \to \varepsilon_n}}{\frac{dN_m}{dt} = A_m N_m}$$

Absorption and induced emission are proportional to $\rho(\nu)$ which is the radiant energy density of the gas as a function of radiant frequency $\nu$. The $A$ and $B$ coefficients are presumed to not depend on temperature $T$.

At equilibrium, transition rates between $\varepsilon_m$ and $\varepsilon_n$ are equal.

$$\underset{\substack{\text{absorption}\\ \varepsilon_n \to \varepsilon_m}}{B_n \rho(\nu) N_n} = \underset{\substack{\text{induced emission}\\ \varepsilon_m \to \varepsilon_n}}{B_m \rho(\nu) N_m} + \underset{\substack{\text{spontaneous emission}\\ \varepsilon_m \to \varepsilon_n}}{A_m N_m}$$

Divide through by $N$ to obtain

$$\underset{\substack{\text{absorption}\\ \varepsilon_n \to \varepsilon_m}}{B_n \rho(\nu) p_n \exp\left(-\frac{\varepsilon_n}{kT}\right)} = \underset{\substack{\text{induced emission}\\ \varepsilon_m \to \varepsilon_n}}{B_m \rho(\nu) p_m \exp\left(-\frac{\varepsilon_m}{kT}\right)} + \underset{\substack{\text{spontaneous emission}\\ \varepsilon_m \to \varepsilon_n}}{A_m p_m \exp\left(-\frac{\varepsilon_m}{kT}\right)}$$

Note that for increasing $T$ we have

$$\lim_{T \to \infty} \exp\left(-\frac{\varepsilon_n}{kT}\right) = 1$$

It follows that for $T \to \infty$ the equilibrium formula is

$$B_n \rho(\nu) p_n = B_m \rho(\nu) p_m + A_m p_m$$

Divide through by $\rho(\nu)$.

$$B_n p_n = B_m p_m + A_m p_m / \rho(\nu)$$

Energy density $\rho(\nu)$ increases with temperature $T$ hence $A_m p_m / \rho(\nu)$ vanishes for $T \to \infty$ leaving

$$B_n p_n = B_m p_m$$

Einstein reasoned that the above relation is true in general based on the assumption that the factors involved do not depend on $T$. By substitution in the absorption term we can now eliminate $B_n p_n$ and obtain

$$\underbrace{B_m \rho(\nu) p_m \exp\left(-\frac{\varepsilon_n}{kT}\right)}_{\substack{\text{absorption} \\ \varepsilon_n \to \varepsilon_m}} = \underbrace{B_m \rho(\nu) p_m \exp\left(-\frac{\varepsilon_m}{kT}\right)}_{\substack{\text{induced emission} \\ \varepsilon_m \to \varepsilon_n}} + \underbrace{A_m p_m \exp\left(-\frac{\varepsilon_m}{kT}\right)}_{\substack{\text{spontaneous emission} \\ \varepsilon_m \to \varepsilon_n}}$$

Divide through by $B_m p_m$ and rearrange terms.

$$\underbrace{\rho(\nu) \exp\left(-\frac{\varepsilon_n}{kT}\right)}_{\substack{\text{absorption} \\ \varepsilon_n \to \varepsilon_m}} - \underbrace{\rho(\nu) \exp\left(-\frac{\varepsilon_m}{kT}\right)}_{\substack{\text{induced emission} \\ \varepsilon_m \to \varepsilon_n}} = \underbrace{(A_m/B_m) \exp\left(-\frac{\varepsilon_m}{kT}\right)}_{\substack{\text{spontaneous emission} \\ \varepsilon_m \to \varepsilon_n}}$$

Solve for energy density $\rho(\nu)$.

$$\rho(\nu) = \frac{(A_m/B_m) \exp\left(-\dfrac{\varepsilon_m}{kT}\right)}{\exp\left(-\dfrac{\varepsilon_n}{kT}\right) - \exp\left(-\dfrac{\varepsilon_m}{kT}\right)} = \frac{A_m/B_m}{\exp\left(\dfrac{\varepsilon_m - \varepsilon_n}{kT}\right) - 1}$$

From Wien's law $\rho(\nu) = \alpha \nu^3 \exp(-h\nu/kT)$ which is accurate for large $\nu$, we have

$$A_m/B_m = \alpha \nu^3$$

and

$$\varepsilon_m - \varepsilon_n = h\nu$$

Planck's law follows directly.

$$\rho(\nu) = \frac{\alpha \nu^3}{\exp\left(\dfrac{h\nu}{kT}\right) - 1}$$