Problem 1

1. Due to the identity between x and y, I only show the result w.r.t. x.

$$\begin{aligned}
F(\nabla_x I) &= \iint \nabla_x I(x,y)\ e^{-2\pi i(\xi x+\eta y)}\ dxdy \\
&= \int I(x,y)\ e^{-2\pi i(\xi x+\eta y)}\ \Big|_{-\infty}^{\infty}\ dy - \iint I(x,y)\ \nabla_x \left[e^{-2\pi i(\xi x+\eta y)}\right]\ dxdy \\
&= 0 - (-2\pi i\xi) \iint I(x,y)\ e^{-2\pi i(\xi x+\eta y)}\ dxdy \\
&= 2\pi i\xi \hat{I}
\end{aligned}$$

The derivations above use $\lim\limits_{x\to\infty} I(x,y) = 0$ and $\hat{I} = \iint I(x,y)\ e^{-2\pi i(\xi x+\eta y)}\ dxdy$.

2. Though $G^*(\xi)$ and $G(\xi)$ are both in the form of Fourier Transform with integrated variable t, these two integrated variable can be regarded independent, and it is equivalent to compose their space together to formulate the product of integrals as a multiple integral. Next by switching the integration order between t and ξ, we can easily eliminate ξ, as shown below:

$$\begin{aligned}
\int G^*(\xi)\ G(\xi)\ d\xi &= \int \left[\int g(t)e^{-2\pi i\xi t}dt\right]^* \left[\int g(t)e^{-2\pi i\xi t}dt\right] d\xi \\
&= \int \left[\iint g^*(t_1)g(t_2)\ e^{2\pi i\xi(t_1-t_2)}dt_1dt_2\right]\ d\xi \\
&= \iint g^*(t_1)g(t_2) \left[\int e^{2\pi i(t_1-t_2)\xi}d\xi\right] dt_1dt_2 \\
&= \iint g^*(t_1)g(t_2)\delta(t_1-t_2)dt_1dt_2 \\
&= \int g^*(t)g(t)\ dt\ = \int |g(t)|^2\, dt
\end{aligned}$$

3. Given $F(\nabla_x I) = 2\pi i\xi\hat{I}$ and $F(\nabla_y I) = 2\pi i\eta\hat{I}$, according to the preservation of energy, we have:

$$\begin{cases}
\beta \iint \left[\nabla_x I\right]^2 dxdy = 4\pi^2\beta \iint \xi^2 \hat{I}^2 d\xi d\eta \\
\beta \iint \left[\nabla_y I\right]^2 dxdy = 4\pi^2\beta \iint \eta^2 \hat{I}^2 d\xi d\eta
\end{cases}$$

So we have: $H(I) = \beta\iint\left[\nabla_x I + \nabla_y I\right]^2 dxdy = 4\pi^2\beta\iint(\xi^2+\eta^2)\hat{I}^2 d\xi d\eta$. On the other hand, the probability $p(I) \propto \exp\left(-H(x)\right)$, which can be factorized to: $p\left(\hat{I}(\xi,\eta)\right) \propto \exp\left(-4\pi^2\beta(\xi^2+\eta^2)\hat{I}^2\right)$. This is the form of Gaussian distribution, where the mean $\mu = 0$ and the variance satisfies $\frac{1}{2}\sigma^{-2} = 4\pi^2\beta(\xi^2+\eta^2)$. Hence variance $\sigma^2 = \frac{1}{8\pi^2\beta(\xi^2+\eta^2)}$.

4. It is obvious that $C = \frac{1}{8\pi^2\beta}$. Next we analytically prove the cumulative power of $A^2(f)$ is constant. Note that in the calculation $\hat{I}(\xi,\eta)$ can be substituted by its mean, regardless of its probabilistic distribution. Let Ω denote the annular region covering the frequency band $[f, 2f]$, we have:

$$\begin{aligned}
\iint_\Omega A^2 \ d\xi d\eta &= \iint_\Omega \frac{C}{\xi^2+\eta^2} \ d\xi d\eta \\
&= \iint_\Omega \frac{C}{f^2} f \ d\theta df \\
&= C \int_f^{2f} df \int_0^{2\pi} \frac{1}{f} \ d\theta \\
&= 2\pi C \int_f^{2f} \frac{df}{f} \\
&= 2\pi C \ln 2
\end{aligned}$$

This invariance is intuitive, since when we scale up or down an image, the range of frequency band should change proportionally. And the cumulative power of corresponding frequency components on the Fourier spectrum should be maintained invariant, which shows the consistency with the calculation above.

Problem 2

1. Given $H(I) = \beta \iint [\nabla_x I]^2 + [\nabla_y I]^2 \, dxdy$ and $\frac{dI}{dt} = -\frac{\delta H}{\delta I}$, we have:

$$\begin{aligned}
\frac{dI}{dt} &= -\frac{\delta H}{\delta I} \\
&= -\beta \left[-\frac{d}{dx} [2\nabla_x I] - \frac{d}{dy} [2\nabla_y I] \right] \\
&= 2\beta(\nabla_{xx} I + \nabla_{yy} I) = 2\beta \Delta I
\end{aligned}$$

Despite the norm, such a dynamic setting can provide the plausible direction for optimization.

2. The Hamiltonian energy can be re-formulated in discrete form as follows:

$$H(I) = \sum_{x,y} I(x,y)^2 + I(x+1,y)^2 + I(x,y+1)^2 - 2\left[I(x+1,y) + I(x,y+1)\right] I(x,y)$$

Here I remind that the relevant terms also include several terms at previous locations $(x-1,y)$ and $(x,y-1)$ when deriving partial derivatives w.r.t. (x,y). The next few steps are as follows:

$$\begin{aligned}
\frac{dI}{dt} &= -\frac{\delta H}{\delta I} \\
&= -\beta\left[8I(x,y) - 2\left(I(x-1,y) + I(x+1,y) + I(x,y-1) + I(x,y+1)\right)\right] \\
&= 2\beta\left[\left(I(x+1,y) + I(x-1,y) - 2I(x,y)\right) + \left(I(x,y+1) + I(x,y-1) - 2I(x,y)\right)\right] \\
&\approx 2\beta[\nabla_{xx}I + \nabla_{yy}I] = 2\beta\Delta I
\end{aligned}$$

3. Actually, the derivation in discrete form has incorporated torus condition, otherwise there will be discrepancy near the edges. To infer the final states of the image is to infer the stationary distribution of H. This also means $\Delta I(x,y) \to 0, \forall x, y$. I conjecture the final image is a totally smooth image with invariant gradients, since this seems to be the most achievable one.

Problem 3

1. Given the condition that, the number of line segments within a unit area is invariant regardless of image scale, we assume a unit A at scale of s covers $\lambda(a,b,A)$ line segments whose length locates in $[a,b]$. If we scale this image to $2s$ and A should expand to $4A$ correspondingly, then we can also observe the λ line segments with doubled length in the area of $4A$. That is, $\lambda(a,b,A) = \lambda(2a,2b,4A)$. According to the uniformity, we can infer $\lambda(2a,2b,4A) = 4\lambda(2a,2b,A)$. To this end, we have $\lambda(a,b,A) = 4\lambda(2a,2b,A)$.

2. Given a fixed A, it is obvious that $\lambda(a,b,A) \propto \int_a^b p(r)dr$. Therefore, from $\lambda(a,b,A) = s^2\lambda(sa,sb,A)$, we can infer $\int_a^b p(r)dr = s^2\int_{sa}^{sb} p(r)dr$.

3. It is easy to verify that $\frac{d\lambda(a,r,A)}{dr} \propto p(r)$. On the other hand, we can formulate it in another way: $\frac{d\lambda(a,r,A)}{dr} = \frac{d}{dr}\left(s^2\lambda(sa,sr,A)\right) \propto s^3p(sr)$, and the proportions are identical. Hence, we derive that $p(r) = s^3p(sr)$. Let $r = 1$, we have: $p(s) = \frac{p(1)}{s^3}$. Obviously, we show that $p(r) \propto r^{-3}$.