1(a). Given an estimate x_k of a local minimizer of $f(x)$, a *linesearch* method (i) computes a search direction s_k, which must also be a descent direction (i.e., $s_k^T \nabla_x f(x_k) < 0$) [1 mark], and (ii) computes a stepsize α_k so that $f(x_k + \alpha_k s_k)$ is "sufficiently" smaller than $f(x_k)$ (using for instance a backtracking Armijo rule) [1 mark]. The next iterate is $x_{k+1} = x_k + \alpha_k s_k$ [2 marks].

By contrast, a *trust region* method computes a trial step s_k to approximately minimize a model approximation $m_k(x_k + s)$ of $f(x_k + s)$ where the step is required to satisfy the trust-region constraint $\|s\| \leq \Delta_k$ for some $\Delta_k > 0$ [1 mark]. If the actual decrease $f(x_k) - f(x_k + s_k)$ is close to that predicted by the model, $m(x_k) - m(x_k + s_k)$, the next iterate is $x_{k+1} = x_k + s_k$, and $\Delta_{k+1} \geq \Delta_k$ [1 mark]. If the actual decrease is significantly worse than that predicted, $x_{k+1} = x_k$, and the new radius $\Delta_{k+1} < \Delta_k$ [1 mark]. The approximate minimizer of the model is required to be at least as good as the Cauchy point [1 mark].

1(b). There exist Lagrange multipliers y_* for which x_* is primal feasible, i.e.,

$$c(x_*) \geq 0, \qquad [\tfrac{1}{2} \text{ mark}]$$

dual feasible i.e.,

$$\nabla_x f(x_*) - (\nabla c(x_*))^T y_* = 0 \text{ and } y_* \geq 0, \qquad [\tfrac{1}{2} \text{ mark}]$$

and satisfies the complementary slackness condition

$$c_i(x_*)(y_*)_i = 0 \text{ for each constraint.} \qquad [1 \text{ mark}]$$

1(c). The gradient of the objective function is $g + Bx$, while the gradient of the constraint $c(x) = \frac{1}{2}\Delta^2 - \frac{1}{2}x^T x \geq 0$ is $-x$ [1 mark]. Thus the dual feasibility equation in 1(b) above gives that

$$g + Bx - (-x_*)\lambda_* = 0 \text{ and } \lambda_* \geq 0$$

i.e., that

$$(B + \lambda_* I)x_* = -g \text{ and } \lambda_* \geq 0. \qquad [1 \text{ mark}]$$

The complementary slackness condition is that

$$(\tfrac{1}{2}\Delta^2 - \tfrac{1}{2}x_*^T x_*)\lambda_* = 0 \qquad [1 \text{ mark}]$$

so that either $\lambda_* = 0$ [1 mark] or $\frac{1}{2}\Delta^2 - \frac{1}{2}x_*^T x_* = 0$ [1 mark]; the second possibility is equivalent to $\|x_*\|_2 = \Delta$.

1(d). The unconstrained minimizer $(-1, 0, -1/2)$ has an ℓ_2-norm of $\sqrt{5}/2 > 5/12$ [1 mark], so the solution must lie on the boundary of the constraint [1 mark]. The solution must be of the form $-(1/(1+\lambda), 0, 1/(2+\lambda))^T$ [1 mark]. To satisfy the trust-region constraint, we then must have

$$\frac{1}{(1+\lambda)^2} + \frac{1}{(2+\lambda)^2} = \Delta^2 = \frac{25}{144}$$

which has a root $\lambda = 2$ [1 mark]. Thus the required solution is $-(1/3, 0, 1/4)^T$[1 mark].

2(a). The first-order optimality conditions are that $x_1 \geq 0$ (primal feasibility) [$\frac{1}{2}$ mark],

$$\begin{pmatrix} 1 \\ x_2 \end{pmatrix} - y \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0$$

and $y \geq 0$ (dual feasibility) [1 mark], and $y \cdot x_1 = 0$ (complementary slackness) [$\frac{1}{2}$ mark]. Dual feasibility says that $y = 1$ and $x_2 = 0$, from which we deduce that $x_1 = 0$ from complementary slackness [1 mark]. Second-order optimality conditions are simply that

$$s_2^2 = (s_1, s_2)^T \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \geq 0$$

for all $s \neq 0$ for which $s_1 = 0$ which are automatically satisfied [1 mark]. Thus the solution is $x = (0, 0)$ with Lagrange multiplier $y = 1$ [1 mark].

2(b). The logarithmic barrier function is

$$\Phi(x, \mu) = x_1 + \tfrac{1}{2} x_2^2 - \mu \log x_1. \qquad \text{[1 mark]}$$

The first-order optimality conditions for the unconstrained minimization of Φ are that

$$\begin{pmatrix} 1 \\ x_2 \end{pmatrix} - \mu \begin{pmatrix} x_1^{-1} \\ 0 \end{pmatrix} = 0. \qquad \text{[1 mark]}$$

If we let $x(\mu)$ be the desired minimizer, the optimality conditions indicate that $x(\mu) = (\mu, 0)$ [1 mark], while the Lagrange multiplier estimates are $y(\mu) = c(x(\mu))/\mu = 1$ [1 mark]. The Hessian is positive definite [1 mark].

2(c). The Hessian matrix of the logarithmic barrier function is

$$\begin{pmatrix} \mu x_1^{-2} & 0 \\ 0 & 1 \end{pmatrix}; \qquad \text{[1 mark]}$$

at the minimizer of $\Phi(x, \mu)$, the Hessian is

$$\begin{pmatrix} \mu^{-1} & 0 \\ 0 & 1 \end{pmatrix}. \qquad \text{[1 mark]}$$

The eigenvalues are 1 and μ^{-1} [1 mark]. As μ goes to zero, one eigenvalue diverges to infinity, while the other one stays fixed at 1 [1 mark]. While this means that the condition number approaches infinity (and thus there *may be* large numerical errors), the growth does not actually happen, since the Newton equations may be reformulated as a well-conditioned system [1 mark].

2(d). The primal-dual system at $x(\mu)$ is

$$\begin{pmatrix} \mu^{-1} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = - \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} - \bar{\mu} \begin{pmatrix} \mu^{-1} \\ 0 \end{pmatrix} \right] \qquad \text{[2 marks]}$$

Thus $s_2 = 0$, while $s_1 = -\mu + \bar{\mu}$ [2 marks]. In particular $x(\mu) + s = \bar{\mu} = x(\bar{\mu})$, the minimizer of $\Phi(x, \bar{\mu})$ [1 mark]