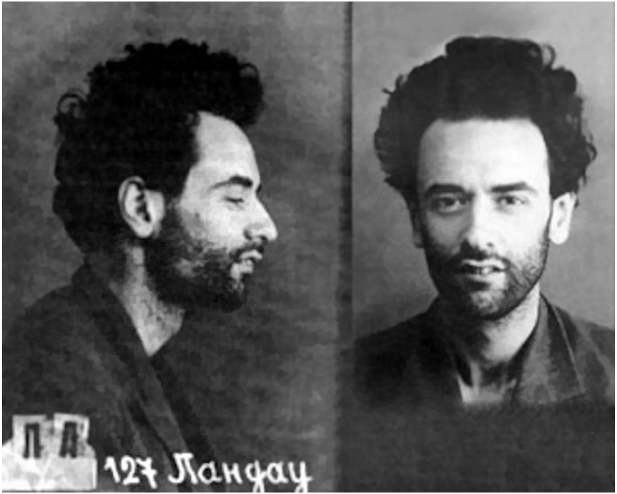
Phase transitions for everybody

Serena Dipierro & Enrico Valdinoci

Swansea 2024

Landau theory of phase transitions



Landau theory of phase transitions

The description of coexisting phases in a given material is a very complex topic, combining different perspectives from statistical physics, thermodynamics, material sciences and mathematical analysis.

Also, the specific case under consideration may require a different setting, given the different underlying physical structures involved: for instance, while in our everyday life we are mostly exposed to the changes of state related to melting, freezing, vaporization and condensation, other important phase transitions, such as the one from a conducting to a superconducting state, are the outcome of brand new properties of the solid state, such as electron coupling.

Landau theory of phase transitions

The description of coexisting phases in a given material is a very complex topic, combining different perspectives from statistical physics, thermodynamics, material sciences and mathematical analysis.

Also, the specific case under consideration may require a different setting, given the different underlying physical structures involved: for instance, while in our everyday life we are mostly exposed to the changes of state related to melting, freezing, vaporization and condensation, other important phase transitions, such as the one from a conducting to a superconducting state, are the outcome of brand new properties of the solid state, such as electron coupling.

Landau theory of phase transitions

A common treat in the study of phase transitions is to describe a given system in terms of relevant physical quantities, such as temperature T, pressure P, volume V, entropy S, magnetic moment M, etc.

The energy of a system is thus an outline of its ability to perform some tasks.

In this performance, however, the system typically wastes some energy in the form of heat.

The “useful energy”, that is the energy that is available to do work, thus consists in the difference between the full internal energy of the system minus the energy that is “unavailable to perform work” since it gets lost through heat.

Being “free to do the work”, such energy is often called free energy, though the name is under an intense debate.

Landau theory of phase transitions

A common treat in the study of phase transitions is to describe a given system in terms of relevant physical quantities, such as temperature T, pressure P, volume V, entropy S, magnetic moment M, etc.

The energy of a system is thus an outline of its ability to perform some tasks.

In this performance, however, the system typically wastes some energy in the form of heat.

The “useful energy”, that is the energy that is available to do work, thus consists in the difference between the full internal energy of the system minus the energy that is “unavailable to perform work” since it gets lost through heat.

Being “free to do the work”, such energy is often called free energy, though the name is under an intense debate.

Landau theory of phase transitions

A common treat in the study of phase transitions is to describe a given system in terms of relevant physical quantities, such as temperature T, pressure P, volume V, entropy S, magnetic moment M, etc.

The energy of a system is thus an outline of its ability to perform some tasks.

In this performance, however, the system typically wastes some energy in the form of heat.

The “useful energy”, that is the energy that is available to do work, thus consists in the difference between the full internal energy of the system minus the energy that is “unavailable to perform work” since it gets lost through heat.

Being “free to do the work”, such energy is often called free energy, though the name is under an intense debate.

Landau theory of phase transitions

A common treat in the study of phase transitions is to describe a given system in terms of relevant physical quantities, such as temperature T, pressure P, volume V, entropy S, magnetic moment M, etc.

The energy of a system is thus an outline of its ability to perform some tasks.

In this performance, however, the system typically wastes some energy in the form of heat.

The "useful energy", that is the energy that is available to do work, thus consists in the difference between the full internal energy of the system minus the energy that is "unavailable to perform work" since it gets lost through heat.

Being "free to do the work", such energy is often called free energy, though the name is under an intense debate.

Landau theory of phase transitions

A common treat in the study of phase transitions is to describe a given system in terms of relevant physical quantities, such as temperature T, pressure P, volume V, entropy S, magnetic moment M, etc.

The energy of a system is thus an outline of its ability to perform some tasks.

In this performance, however, the system typically wastes some energy in the form of heat.

The “useful energy”, that is the energy that is available to do work, thus consists in the difference between the full internal energy of the system minus the energy that is “unavailable to perform work” since it gets lost through heat.

Being “free to do the work”, such energy is often called free energy, though the name is under an intense debate.

Landau theory of phase transitions

Hence, the minimizers (or, more generally, the stable critical points) of this energy correspond to observable states of our system.

The system may present significantly different features, or “phases”, such as being in a solid or fluid state, or having a magnetic momentum, or presenting superfluid or superconductor properties, and the appearance of these phases may be seen as an outcome of energy minimization.

Landau theory of phase transitions

Hence, the minimizers (or, more generally, the stable critical points) of this energy correspond to observable states of our system.

The system may present significantly different features, or “phases”, such as being in a solid or fluid state, or having a magnetic momentum, or presenting superfluid or superconductor properties, and the appearance of these phases may be seen as an outcome of energy minimization.

Landau theory of phase transitions

The arising of different phases may be the outcome of a critical physical parameter involved in the free energy, such as temperature: a “disordered phase” typically corresponds to a high temperature, while an “exceptionally ordered phase” arises at low temperatures.

This is the case, for instance, for magnetization, since magnetic materials have no permanent magnetic moment above their Curie Temperature (about 570 degree Celsius for the usual magnetite) but below this temperature the atoms tend to behave as tiny magnets which spontaneously align themselves, so that the magnetic materials show a permanent magnetization oriented in a certain direction.

Landau theory of phase transitions

The arising of different phases may be the outcome of a critical physical parameter involved in the free energy, such as temperature: a “disordered phase” typically corresponds to a high temperature, while an “exceptionally ordered phase” arises at low temperatures.

This is the case, for instance, for magnetization, since magnetic materials have no permanent magnetic moment above their Curie Temperature (about 570 degree Celsius for the usual magnetite) but below this temperature the atoms tend to behave as tiny magnets which spontaneously align themselves, so that the magnetic materials show a permanent magnetization oriented in a certain direction.

Landau theory of phase transitions

The arising of different phases may be the outcome of a critical physical parameter involved in the free energy, such as temperature: a "disordered phase" typically corresponds to a high temperature, while an "exceptionally ordered phase" arises at low temperatures.

This is the case, for instance, for magnetization, since magnetic materials have no permanent magnetic moment above their Curie Temperature (about 570 degree Celsius for the usual magnetite) but below this temperature the atoms tend to behave as tiny magnets which spontaneously align themselves, so that the magnetic materials show a permanent magnetization oriented in a certain direction.

Landau theory of phase transitions

When the temperature T is above the Curie Temperature T_c, in the absence of external sources a magnetic system lies in a zero field state, which happens to be a minimal configuration for the free energy corresponding to its temperature.

When the temperature is decreased below such critical temperature T_c, the system will go through a state in which the magnetization is still zero, but this corresponds only to a critical point, not a minimum of the free energy, making this equilibrium configuration totally unstable.

Below the critical temperature T_c, small perturbations from the environment will inevitably induce the system to reorganize its microscopic structure to preserve a zero average but creating regions with a nontrivial magnetic momentum.

Landau theory of phase transitions

When the temperature T is above the Curie Temperature T_c, in the absence of external sources a magnetic system lies in a zero field state, which happens to be a minimal configuration for the free energy corresponding to its temperature.

When the temperature is decreased below such critical temperature T_c, the system will go through a state in which the magnetization is still zero, but this corresponds only to a critical point, not a minimum of the free energy, making this equilibrium configuration totally unstable.

Below the critical temperature T_c, small perturbations from the environment will inevitably induce the system to reorganize its microscopic structure to preserve a zero average but creating regions with a nontrivial magnetic momentum.

Landau theory of phase transitions

When the temperature T is above the Curie Temperature T_c, in the absence of external sources a magnetic system lies in a zero field state, which happens to be a minimal configuration for the free energy corresponding to its temperature.

When the temperature is decreased below such critical temperature T_c, the system will go through a state in which the magnetization is still zero, but this corresponds only to a critical point, not a minimum of the free energy, making this equilibrium configuration totally unstable.

Below the critical temperature T_c, small perturbations from the environment will inevitably induce the system to reorganize its microscopic structure to preserve a zero average but creating regions with a nontrivial magnetic momentum.

Landau theory of phase transitions

The formation of these magnetic domains will produce a supplementary interfacial energy, which, in some sense, favors domain segregation with a phase separation which is “as small as possible”.

Such a phase separation also produces a symmetry breaking: the free energy is symmetric (since it weighs equally, say, the magnetizations oriented towards the North pole and the ones oriented towards the South pole), nonetheless the magnetization configuration reached by the system during the cooling is somewhat *accidental*, as a result of small environmental perturbations, making the final state reached by the system *not necessarily symmetric*.

Landau theory of phase transitions

The formation of these magnetic domains will produce a supplementary interfacial energy, which, in some sense, favors domain segregation with a phase separation which is “as small as possible”.

Such a phase separation also produces a symmetry breaking: the free energy is symmetric (since it weighs equally, say, the magnetizations oriented towards the North pole and the ones oriented towards the South pole), nonetheless the magnetization configuration reached by the system during the cooling is somewhat *accidental*, as a result of small environmental perturbations, making the final state reached by the system *not necessarily symmetric*.

Landau theory of phase transitions

The formation of these magnetic domains will produce a supplementary interfacial energy, which, in some sense, favors domain segregation with a phase separation which is “as small as possible”.

Such a phase separation also produces a symmetry breaking: the free energy is symmetric (since it weighs equally, say, the magnetizations oriented towards the North pole and the ones oriented towards the South pole), nonetheless the magnetization configuration reached by the system during the cooling is somewhat *accidental*, as a result of small environmental perturbations, making the final state reached by the system *not necessarily symmetric*.

Landau theory of phase transitions

To account for the phenomena of phase transition and phase coexistence, one can consider an order parameter η which describes *how every point of the system is "organized" with respect to a given notion of phase*.

The order parameter can be either a scalar or a vector: for instance, in a liquid-gas phase transition the order parameter corresponds to the difference of the densities between the two phases, which is a scalar, while in superfluidity and superconductivity it is a complex-valued wave function (or, equivalently, a two-dimensional vector), and for ferromagnetic momenta it is in general a three-dimensional vector.

Phase transitions also occur in cosmology, since as the universe expanded and cooled, a number of symmetry-breaking phase transitions occurred, and the description of these phenomena often relies on an order parameter which is a tensor.

Here we only consider the case in which the order parameter η is a scalar function.

Landau theory of phase transitions

To account for the phenomena of phase transition and phase coexistence, one can consider an order parameter η which describes *how every point of the system is "organized" with respect to a given notion of phase*.

The order parameter can be either a scalar or a vector: for instance, in a liquid-gas phase transition the order parameter corresponds to the difference of the densities between the two phases, which is a scalar, while in superfluidity and superconductivity it is a complex-valued wave function (or, equivalently, a two-dimensional vector), and for ferromagnetic momenta it is in general a three-dimensional vector.

Phase transitions also occur in cosmology, since as the universe expanded and cooled, a number of symmetry-breaking phase transitions occurred, and the description of these phenomena often relies on an order parameter which is a tensor.

Here we only consider the case in which the order parameter η is a scalar function.

Landau theory of phase transitions

To account for the phenomena of phase transition and phase coexistence, one can consider an order parameter η which describes *how every point of the system is "organized" with respect to a given notion of phase*.

The order parameter can be either a scalar or a vector: for instance, in a liquid-gas phase transition the order parameter corresponds to the difference of the densities between the two phases, which is a scalar, while in superfluidity and superconductivity it is a complex-valued wave function (or, equivalently, a two-dimensional vector), and for ferromagnetic momenta it is in general a three-dimensional vector.

Phase transitions also occur in cosmology, since as the universe expanded and cooled, a number of symmetry-breaking phase transitions occurred, and the description of these phenomena often relies on an order parameter which is a tensor.

Here we only consider the case in which the order parameter η is a scalar function.

Landau theory of phase transitions

To account for the phenomena of phase transition and phase coexistence, one can consider an order parameter η which describes *how every point of the system is "organized" with respect to a given notion of phase*.

The order parameter can be either a scalar or a vector: for instance, in a liquid-gas phase transition the order parameter corresponds to the difference of the densities between the two phases, which is a scalar, while in superfluidity and superconductivity it is a complex-valued wave function (or, equivalently, a two-dimensional vector), and for ferromagnetic momenta it is in general a three-dimensional vector.

Phase transitions also occur in cosmology, since as the universe expanded and cooled, a number of symmetry-breaking phase transitions occurred, and the description of these phenomena often relies on an order parameter which is a tensor.

Here we only consider the case in which the order parameter η is a scalar function.

The second-order theory of phase transitions

Consider a critical situation in which phase separation occurs. Assume that the free energy presents the Taylor expansion

$$a_0 + a_1\eta + a_2\eta^2 + a_3\eta^3 + a_4\eta^4 + \dots,$$

where the coefficients a_0, a_1, a_2, a_3, a_4, …, depend on relevant physical quantities, say the temperature T of the system.

Suppose the free energy is symmetric with respect to the order parameter (say, assuming that the deviations from the neutral case equally affect the energy): then, the odd coefficients must vanish, thus reducing the free energy (up to higher orders) to

$$a_0 + a_2\eta^2 + a_4\eta^4.$$

The second-order theory of phase transitions

Consider a critical situation in which phase separation occurs. Assume that the free energy presents the Taylor expansion

$$a_0 + a_1\eta + a_2\eta^2 + a_3\eta^3 + a_4\eta^4 + \dots,$$

where the coefficients a_0, a_1, a_2, a_3, a_4, $\dots$, depend on relevant physical quantities, say the temperature T of the system.

Suppose the free energy is symmetric with respect to the order parameter (say, assuming that the deviations from the neutral case equally affect the energy): then, the odd coefficients must vanish, thus reducing the free energy (up to higher orders) to

$$a_0 + a_2\eta^2 + a_4\eta^4.$$

The second-order theory of phase transitions

Consider a critical situation in which phase separation occurs. Assume that the free energy presents the Taylor expansion

$$a_0 + a_1\eta + a_2\eta^2 + a_3\eta^3 + a_4\eta^4 + \dots,$$

where the coefficients $a_0, a_1, a_2, a_3, a_4, \dots,$ depend on relevant physical quantities, say the temperature T of the system.

Suppose the free energy is symmetric with respect to the order parameter (say, assuming that the deviations from the neutral case equally affect the energy): then, the odd coefficients must vanish, thus reducing the free energy (up to higher orders) to

$$a_0 + a_2\eta^2 + a_4\eta^4.$$

The second-order theory of phase transitions

The coefficient a_0 is irrelevant to determine the critical points, but the coefficients a_2 and a_4 play an essential role. So, we can think that a_0 is just a constant, while $a_2 = a_2(T)$ and $a_4 = a_4(T)$ depend on the temperature T.

If we expect the state parameter η to be confined in a bounded region (which is typically the case, since we do not expect that a physical parameter tends to diverge), it is convenient to assume that $a_4(T)$ is positive for all T (in this way, the free energy is bounded from below and possesses minima for all T).

To model the spontaneous formations of new phases below the critical temperature, one may suppose that

$$a_2(T) > 0 \text{ for all } T > T_c \text{ and } a_2(T) < 0 \text{ for all } T < T_c.$$

The second-order theory of phase transitions

The coefficient a_0 is irrelevant to determine the critical points, but the coefficients a_2 and a_4 play an essential role. So, we can think that a_0 is just a constant, while $a_2 = a_2(T)$ and $a_4 = a_4(T)$ depend on the temperature T.

If we expect the state parameter η to be confined in a bounded region (which is typically the case, since we do not expect that a physical parameter tends to diverge), it is convenient to assume that $a_4(T)$ is positive for all T (in this way, the free energy is bounded from below and possesses minima for all T).

To model the spontaneous formations of new phases below the critical temperature, one may suppose that

$$a_2(T) > 0 \text{ for all } T > T_c \text{ and } a_2(T) < 0 \text{ for all } T < T_c.$$

The second-order theory of phase transitions

The coefficient a_0 is irrelevant to determine the critical points, but the coefficients a_2 and a_4 play an essential role. So, we can think that a_0 is just a constant, while $a_2 = a_2(T)$ and $a_4 = a_4(T)$ depend on the temperature T.

If we expect the state parameter η to be confined in a bounded region (which is typically the case, since we do not expect that a physical parameter tends to diverge), it is convenient to assume that $a_4(T)$ is positive for all T (in this way, the free energy is bounded from below and possesses minima for all T).

To model the spontaneous formations of new phases below the critical temperature, one may suppose that

$$a_2(T) > 0 \text{ for all } T > T_c \text{ and } a_2(T) < 0 \text{ for all } T < T_c.$$

The second-order theory of phase transitions

Assuming that a_2 varies continuously with respect to T,

$$a_2(T_c) = 0.$$

In this way, one readily checks that the critical points are

$$\begin{cases} \{0\} & \text{if } T \geq T_c, \\ \left\{ -\sqrt{-\dfrac{a_2(T)}{2a_4(T)}},\ 0,\ \sqrt{-\dfrac{a_2(T)}{2a_4(T)}} \right\} & \text{if } T < T_c. \end{cases}$$

The second-order theory of phase transitions

Assuming that a_2 varies continuously with respect to T,

$$a_2(T_c) = 0.$$

In this way, one readily checks that the critical points are

$$\begin{cases} \{0\} & \text{if } T \geq T_c, \\ \left\{-\sqrt{-\dfrac{a_2(T)}{2a_4(T)}},\ 0,\ \sqrt{-\dfrac{a_2(T)}{2a_4(T)}}\right\} & \text{if } T < T_c. \end{cases}$$

The second-order theory of phase transitions

Furthermore,

the critical point 0 is a nondegenerate minimum when $T > T_c$,

a degenerate minimum when $T = T_c$,

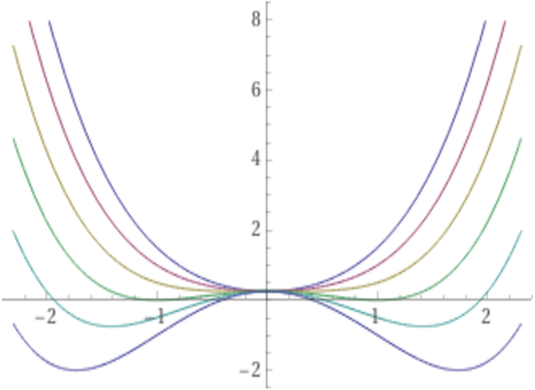
and a local maximum when $T < T_c$,

while $\pm\sqrt{-\frac{a_2(T)}{2a_4(T)}}$ are nondegenerate minima when $T < T_c$.

The second-order theory of phase transitions

As an example, suppose

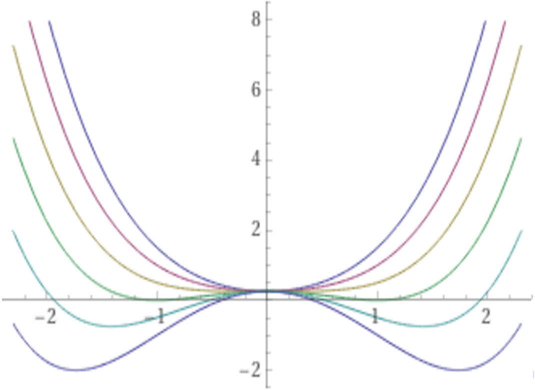
$$T_c := 2, \qquad a_0 := \frac{1}{4}, \qquad a_2(T) := \frac{T-2}{2} \qquad \text{and} \qquad a_4(T) := \frac{1}{4}.$$



The second-order theory of phase transitions

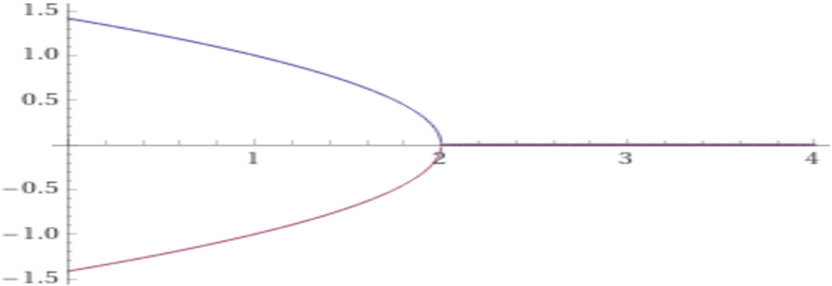
The minimizers of the free energy in this case are

$$\begin{cases} \{0\} & \text{if } T \geq 2, \\ \left\{-\sqrt{2-T},\ \sqrt{2-T}\right\} & \text{if } T < 2, \end{cases} \tag{1}$$



The second-order theory of phase transitions

The bifurcation diagram for the minimizers is therefore



The second-order theory of phase transitions

This bifurcation diagram represents a situation in which the order parameter exhibited by the system changes continuously with respect to the temperature.

Namely, if the free energy depends continuously on the temperature T, then the new minima when $T < T_c$ can be seen as a continuous modification of the null state, since they are given by $\pm\sqrt{2 - T}$ for $T < 2$.

Interestingly, the dependence of these minima on the temperature T is not in general smooth, due to the presence of the square root.

In jargon, the situations in which the observed order parameter depends continuously on the temperature are called second-order phase transitions.

The second-order theory of phase transitions

This bifurcation diagram represents a situation in which the order parameter exhibited by the system changes continuously with respect to the temperature.

Namely, if the free energy depends continuously on the temperature T, then the new minima when $T < T_c$ can be seen as a continuous modification of the null state, since they are given by $\pm\sqrt{2-T}$ for $T < 2$.

Interestingly, the dependence of these minima on the temperature T is not in general smooth, due to the presence of the square root.

In jargon, the situations in which the observed order parameter depends continuously on the temperature are called second-order phase transitions.

The second-order theory of phase transitions

This bifurcation diagram represents a situation in which the order parameter exhibited by the system changes continuously with respect to the temperature.

Namely, if the free energy depends continuously on the temperature T, then the new minima when $T < T_c$ can be seen as a continuous modification of the null state, since they are given by $\pm\sqrt{2-T}$ for $T < 2$.

Interestingly, the dependence of these minima on the temperature T is not in general smooth, due to the presence of the square root.

In jargon, the situations in which the observed order parameter depends continuously on the temperature are called second-order phase transitions.

The second-order theory of phase transitions

This bifurcation diagram represents a situation in which the order parameter exhibited by the system changes continuously with respect to the temperature.

Namely, if the free energy depends continuously on the temperature T, then the new minima when $T < T_c$ can be seen as a continuous modification of the null state, since they are given by $\pm\sqrt{2 - T}$ for $T < 2$.

Interestingly, the dependence of these minima on the temperature T is not in general smooth, due to the presence of the square root.

In jargon, the situations in which the observed order parameter depends continuously on the temperature are called second-order phase transitions.

The second-order theory of phase transitions

The name possibly comes from this: if we evaluate the free energy at the minima, we obtain the free energy as a function of temperature

$$\begin{aligned}
\mathcal{E}(T) \quad &:= \begin{cases} a_0 & \text{if } T \geq T_c, \\ a_0 - \dfrac{a_2^2(T)}{2a_4(T)} + \dfrac{a_4(T)\, a_2^2(T)}{4a_4^2(T)} & \text{if } T < T_c, \end{cases} \\
&= \begin{cases} a_0 & \text{if } T \geq T_c, \\ a_0 - \dfrac{a_2^2(T)}{4a_4(T)} & \text{if } T < T_c. \end{cases}
\end{aligned}$$

The second-order theory of phase transitions

Moreover,

$$\mathcal{E}'(T) = \begin{cases} 0 & \text{if } T > T_c, \\ -\dfrac{a_2(T)\,a_2'(T)}{2a_4(T)} + \dfrac{a_2^2(T)\,a_4'(T)}{4a_4^2(T)} & \text{if } T < T_c, \end{cases}$$

whence

$$\lim_{T\nearrow T_c} \mathcal{E}'(T) = 0 = \lim_{T\searrow T_c} \mathcal{E}'(T).$$

Accordingly, the first derivative with respect to temperature of the "free energy as a function of temperature" vanishes continuously at the critical temperature.

The second-order theory of phase transitions

Moreover,

$$\mathcal{E}'(T) = \begin{cases} 0 & \text{if } T > T_c, \\ -\dfrac{a_2(T)\,a_2'(T)}{2a_4(T)} + \dfrac{a_2^2(T)\,a_4'(T)}{4a_4^2(T)} & \text{if } T < T_c, \end{cases}$$

whence

$$\lim_{T \nearrow T_c} \mathcal{E}'(T) = 0 = \lim_{T \searrow T_c} \mathcal{E}'(T).$$

Accordingly, the first derivative with respect to temperature of the "free energy as a function of temperature" vanishes continuously at the critical temperature.

The second-order theory of phase transitions

Moreover,

$$\mathcal{E}'(T) = \begin{cases} 0 & \text{if } T > T_c, \\ -\dfrac{a_2(T)\, a_2'(T)}{2a_4(T)} + \dfrac{a_2^2(T)\, a_4'(T)}{4a_4^2(T)} & \text{if } T < T_c, \end{cases}$$

whence

$$\lim_{T \nearrow T_c} \mathcal{E}'(T) = 0 = \lim_{T \searrow T_c} \mathcal{E}'(T).$$

Accordingly, the first derivative with respect to temperature of the “free energy as a function of temperature” vanishes continuously at the critical temperature.

The second-order theory of phase transitions

Furthermore,

$$\mathcal{E}''(T) = \begin{cases} 0 & \text{if } T > T_c, \\ -\dfrac{(a_2'(T))^2}{2a_4(T)} - \dfrac{a_2(T)\,a_2''(T)}{2a_4(T)} + \dfrac{a_2(T)\,a_2'(T)\,a_4'(T)}{a_4^2(T)} & \\ \qquad + \dfrac{a_2^2(T)\,a_4''(T)}{4a_4^2(T)} - \dfrac{a_2^2(T)\,(a_4'(T))^2}{2a_4^3(T)} & \text{if } T < T_c, \end{cases}$$

leading to

$$\lim_{T\nearrow T_c} \mathcal{E}''(T) = -\frac{(a_2'(T_c))^2}{2a_4(T_c)} \qquad \text{while} \qquad \lim_{T\searrow T_c} \mathcal{E}''(T) = 0.$$

In particular, if T_c is a nondegenerate zero of a_2 (as it happens for instance in the model case),

$$\lim_{T\nearrow T_c} \mathcal{E}''(T) < 0 = \lim_{T\searrow T_c} \mathcal{E}''(T).$$

The second-order theory of phase transitions

Furthermore,

$$\mathcal{E}''(T) = \begin{cases} 0 & \text{if } T > T_c, \\ -\dfrac{(a_2'(T))^2}{2a_4(T)} - \dfrac{a_2(T)\,a_2''(T)}{2a_4(T)} + \dfrac{a_2(T)\,a_2'(T)\,a_4'(T)}{a_4^2(T)} \\ \qquad + \dfrac{a_2^2(T)\,a_4''(T)}{4a_4^2(T)} - \dfrac{a_2^2(T)\,(a_4'(T))^2}{2a_4^3(T)} & \text{if } T < T_c, \end{cases}$$

leading to

$$\lim_{T \nearrow T_c} \mathcal{E}''(T) = -\frac{(a_2'(T_c))^2}{2a_4(T_c)} \quad \text{while} \quad \lim_{T \searrow T_c} \mathcal{E}''(T) = 0.$$

In particular, if T_c is a nondegenerate zero of a_2 (as it happens for instance in the model case),

$$\lim_{T \nearrow T_c} \mathcal{E}''(T) < 0 = \lim_{T \searrow T_c} \mathcal{E}''(T).$$

The second-order theory of phase transitions

Furthermore,

$$\mathcal{E}''(T) = \begin{cases} 0 & \text{if } T > T_c, \\ -\dfrac{(a_2'(T))^2}{2a_4(T)} - \dfrac{a_2(T)\,a_2''(T)}{2a_4(T)} + \dfrac{a_2(T)\,a_2'(T)\,a_4'(T)}{a_4^2(T)} \\ \qquad + \dfrac{a_2^2(T)\,a_4''(T)}{4a_4^2(T)} - \dfrac{a_2^2(T)\,(a_4'(T))^2}{2a_4^3(T)} & \text{if } T < T_c, \end{cases}$$

leading to

$$\lim_{T \nearrow T_c} \mathcal{E}''(T) = -\frac{(a_2'(T_c))^2}{2a_4(T_c)} \quad \text{while} \quad \lim_{T \searrow T_c} \mathcal{E}''(T) = 0.$$

In particular, if T_c is a nondegenerate zero of a_2 (as it happens for instance in the model case),

$$\lim_{T \nearrow T_c} \mathcal{E}''(T) < 0 = \lim_{T \searrow T_c} \mathcal{E}''(T).$$

The second-order theory of phase transitions

Thus the second derivative with respect to temperature of the “free energy as a function of temperature” presents a discontinuity at the critical temperature.

Maybe, this is the justification of the name of “second-order phase transitions”...

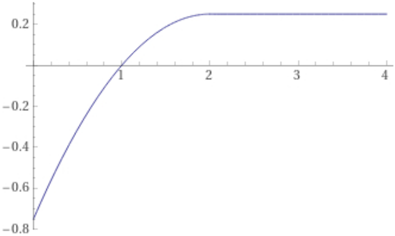
The second-order theory of phase transitions

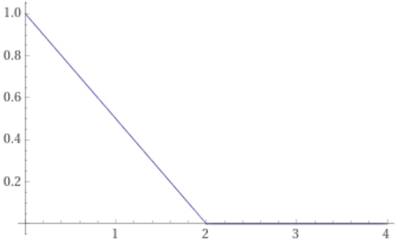
Thus the second derivative with respect to temperature of the “free energy as a function of temperature” presents a discontinuity at the critical temperature.

Maybe, this is the justification of the name of “second-order phase transitions”...

The second-order theory of phase transitions

Here is a sketch of the functions $\mathcal{E}$ and $\mathcal{E}'$ in the model case:





The second-order theory of phase transitions

From a physical point of view, the discontinuities of $\mathcal{E}'$ at the critical temperature are related to the latent heat (roughly speaking, the energy released or absorbed by the system in a phase change without changing its temperature), hence the second-order phase transitions correspond to the absence of latent heat.

The second-order theory of phase transitions

From a physical point of view, the discontinuities of $\mathcal{E}'$ at the critical temperature are related to the latent heat (roughly speaking, the energy released or absorbed by the system in a phase change without changing its temperature), hence the second-order phase transitions correspond to the absence of latent heat.

The second-order theory of phase transitions

A heuristic explanation of the link between the latent heat and the possible discontinuities of $\mathcal{E}'$ can be obtained as follows.

Consider a system undergoing a phase change at temperature T_c, keeping the other physical parameters constant. The derivative with respect to temperature of the free energy corresponds, up to a sign change, to entropy, hence

$$S = -\mathcal{E}' .$$

Also, by the Second Law of Thermodynamics,

$$dS = \frac{dQ}{T} .$$

In this way, we formally have that

$$dQ = T\, dS = -T\, d\mathcal{E}' = -T\, \mathcal{E}''\, dT .$$

The second-order theory of phase transitions

A heuristic explanation of the link between the latent heat and the possible discontinuities of $\mathcal{E}'$ can be obtained as follows.

Consider a system undergoing a phase change at temperature T_c, keeping the other physical parameters constant. The derivative with respect to temperature of the free energy corresponds, up to a sign change, to entropy, hence

$$S = -\mathcal{E}'.$$

Also, by the Second Law of Thermodynamics,

$$dS = \frac{dQ}{T}.$$

In this way, we formally have that

$$dQ = T\,dS = -T\,d\mathcal{E}' = -T\,\mathcal{E}''\,dT.$$

The second-order theory of phase transitions

A heuristic explanation of the link between the latent heat and the possible discontinuities of $\mathcal{E}'$ can be obtained as follows.

Consider a system undergoing a phase change at temperature T_c, keeping the other physical parameters constant. The derivative with respect to temperature of the free energy corresponds, up to a sign change, to entropy, hence

$$S = -\mathcal{E}'.$$

Also, by the Second Law of Thermodynamics,

$$dS = \frac{dQ}{T}.$$

In this way, we formally have that

$$dQ = T\,dS = -T\,d\mathcal{E}' = -T\,\mathcal{E}''\,dT.$$

The second-order theory of phase transitions

A heuristic explanation of the link between the latent heat and the possible discontinuities of $\mathcal{E}'$ can be obtained as follows.

Consider a system undergoing a phase change at temperature T_c, keeping the other physical parameters constant. The derivative with respect to temperature of the free energy corresponds, up to a sign change, to entropy, hence

$$S = -\mathcal{E}'.$$

Also, by the Second Law of Thermodynamics,

$$dS = \frac{dQ}{T}.$$

In this way, we formally have that

$$dQ = T\,dS = -T\,d\mathcal{E}' = -T\,\mathcal{E}''\,dT.$$

The second-order theory of phase transitions

But we need to interpret this equation with a pinch of salt, since $d\mathcal{E}' =$ may be not classically defined at T_c (due to the possible discontinuities of $\mathcal{E}'$ at the critical temperature).

Therefore, possibly integrating a Dirac's Delta,

$$Q(T_c + \varepsilon) - Q(T_c - \varepsilon) = T_c\Big(\mathcal{E}'(T_c + \varepsilon) - \mathcal{E}'(T_c - \varepsilon)\Big).$$

This shows that a discontinuity of the derivative of the free energy at the critical temperature corresponds to a latent heat proportional to such a discontinuity times the critical temperature.

The second-order theory of phase transitions

But we need to interpret this equation with a pinch of salt, since $d\mathcal{E}' =$ may be not classically defined at T_c (due to the possible discontinuities of $\mathcal{E}'$ at the critical temperature).

Therefore, possibly integrating a Dirac's Delta,

$$Q(T_c + \varepsilon) - Q(T_c - \varepsilon) = T_c \Big(\mathcal{E}'(T_c + \varepsilon) - \mathcal{E}'(T_c - \varepsilon) \Big).$$

This shows that a discontinuity of the derivative of the free energy at the critical temperature corresponds to a latent heat proportional to such a discontinuity times the critical temperature.

The second-order theory of phase transitions

But we need to interpret this equation with a pinch of salt, since $d\mathcal{E}' =$ may be not classically defined at T_c (due to the possible discontinuities of $\mathcal{E}'$ at the critical temperature).

Therefore, possibly integrating a Dirac's Delta,

$$Q(T_c + \varepsilon) - Q(T_c - \varepsilon) = T_c \Big(\mathcal{E}'(T_c + \varepsilon) - \mathcal{E}'(T_c - \varepsilon) \Big).$$

This shows that a discontinuity of the derivative of the free energy at the critical temperature corresponds to a latent heat proportional to such a discontinuity times the critical temperature.

The first-order theory of phase transitions

In many physical situations, however, the change of the state of a substance at its critical temperature is related to a latent heat which is supplied to or extracted from the medium without changing its temperature.

These types of phase transitions correspond to a discontinuity of the first derivative of $\mathcal{E}$ and are called first-order phase transitions.

In these situations, the observed order parameters also jump discontinuously at the transition temperature.

To describe these phenomena, we retake the free energy expansion and we aim at modeling a situation in which $\eta = 0$ is the observed value of the state parameter above a critical temperature T_c, but below T_c a new stable phase arises.

The first-order theory of phase transitions

In many physical situations, however, the change of the state of a substance at its critical temperature is related to a latent heat which is supplied to or extracted from the medium without changing its temperature.

These types of phase transitions correspond to a discontinuity of the first derivative of $\mathcal{E}$ and are called first-order phase transitions.

In these situations, the observed order parameters also jump discontinuously at the transition temperature.

To describe these phenomena, we retake the free energy expansion and we aim at modeling a situation in which $\eta = 0$ is the observed value of the state parameter above a critical temperature T_c, but below T_c a new stable phase arises.

The first-order theory of phase transitions

In many physical situations, however, the change of the state of a substance at its critical temperature is related to a latent heat which is supplied to or extracted from the medium without changing its temperature.

These types of phase transitions correspond to a discontinuity of the first derivative of $\mathcal{E}$ and are called first-order phase transitions.

In these situations, the observed order parameters also jump discontinuously at the transition temperature.

To describe these phenomena, we retake the free energy expansion and we aim at modeling a situation in which $\eta = 0$ is the observed value of the state parameter above a critical temperature T_c, but below T_c a new stable phase arises.

The first-order theory of phase transitions

In many physical situations, however, the change of the state of a substance at its critical temperature is related to a latent heat which is supplied to or extracted from the medium without changing its temperature.

These types of phase transitions correspond to a discontinuity of the first derivative of $\mathcal{E}$ and are called first-order phase transitions.

In these situations, the observed order parameters also jump discontinuously at the transition temperature.

To describe these phenomena, we retake the free energy expansion and we aim at modeling a situation in which $\eta = 0$ is the observed value of the state parameter above a critical temperature T_c, but below T_c a new stable phase arises.

The first-order theory of phase transitions

We describe a model in which $\eta = 0$ is a nondegenerate local minimum, hence a stable phase, for the free energy for all values of the temperature T (and the only minimizer when $T > T_c$), but a new stable phase arises when $T \leq T_c$, with the new phase becoming a global minimizer when $T < T_c$.

Here, we are not assuming that the free energy is symmetric in η. The condition that $\eta = 0$ is a critical point gives that a_1 must necessarily vanish for all T, therefore the free energy is

$$a_2(T)\,\eta^2 + a_3(T)\eta^3 + a_4(T)\eta^4.$$

The first-order theory of phase transitions

We describe a model in which $\eta = 0$ is a nondegenerate local minimum, hence a stable phase, for the free energy for all values of the temperature T (and the only minimizer when $T > T_c$), but a new stable phase arises when $T \leq T_c$, with the new phase becoming a global minimizer when $T < T_c$.

Here, we are not assuming that the free energy is symmetric in η. The condition that $\eta = 0$ is a critical point gives that a_1 must necessarily vanish for all T, therefore the free energy is

$$a_2(T)\,\eta^2 + a_3(T)\eta^3 + a_4(T)\eta^4.$$

The first-order theory of phase transitions

We describe a model in which $\eta = 0$ is a nondegenerate local minimum, hence a stable phase, for the free energy for all values of the temperature T (and the only minimizer when $T > T_c$), but a new stable phase arises when $T \leq T_c$, with the new phase becoming a global minimizer when $T < T_c$.

Here, we are not assuming that the free energy is symmetric in η. The condition that $\eta = 0$ is a critical point gives that a_1 must necessarily vanish for all T, therefore the free energy is

$$a_2(T)\,\eta^2 + a_3(T)\eta^3 + a_4(T)\eta^4.$$

The first-order theory of phase transitions

The condition that $\eta = 0$ is a nondegenerate local minimum yields that

$$a_2(T) > 0.$$

The condition that the energy is bounded from below (thus producing minimizers) also gives that

$$a_4(T) > 0.$$

The phase transition can be thus modeled on the specific properties of $a_3(T)$. Namely, the assumption that $\eta = 0$ is the only minimizer for $T > T_c$ says that

$$a_2(T) + a_3(T)\eta + a_4(T)\eta^2 > 0 \qquad \text{for all } \eta \in \mathbb{R} \text{ and } T > T_c.$$

The first-order theory of phase transitions

The condition that $\eta = 0$ is a nondegenerate local minimum yields that

$$a_2(T) > 0.$$

The condition that the energy is bounded from below (thus producing minimizers) also gives that

$$a_4(T) > 0.$$

The phase transition can be thus modeled on the specific properties of $a_3(T)$. Namely, the assumption that $\eta = 0$ is the only minimizer for $T > T_c$ says that

$$a_2(T) + a_3(T)\eta + a_4(T)\eta^2 > 0 \qquad \text{for all } \eta \in \mathbb{R} \text{ and } T > T_c.$$

The first-order theory of phase transitions

The condition that $\eta = 0$ is a nondegenerate local minimum yields that

$$a_2(T) > 0.$$

The condition that the energy is bounded from below (thus producing minimizers) also gives that

$$a_4(T) > 0.$$

The phase transition can be thus modeled on the specific properties of $a_3(T)$. Namely, the assumption that $\eta = 0$ is the only minimizer for $T > T_c$ says that

$$a_2(T) + a_3(T)\eta + a_4(T)\eta^2 > 0 \qquad \text{for all } \eta \in \mathbb{R} \text{ and } T > T_c.$$

The first-order theory of phase transitions

Also, we assume that at $T = T_c$ a new minimizer, say at $\eta = \eta_\star > 0$, occurs, whence

$$a_2(T_c)\,\eta_\star^2 + a_3(T_c)\eta_\star^3 + a_4(T_c)\eta_\star^4 = 0.$$

The existence of a global minimum different from $\eta = 0$ below the critical temperature translates into

$$\min_{\eta\in\mathbb{R}} a_2(T)\,\eta^2 + a_3(T)\eta^3 + a_4(T)\eta^4 < 0 \qquad \text{for all } T < T_c.$$

An example of coefficients satisfying all these conditions is, for instance:

$T_c := 2, \qquad a_2(T) := 1,$

a_3 a smooth and increasing function such that $a_3(T) < 2$ for all $T \in \mathbb{R}$

with $a_3(T) = T - 4$ for all $T \leq 5$,

$a_4(T) := 1 \qquad$ and $\qquad \eta_\star := 1.$

The first-order theory of phase transitions

Also, we assume that at $T = T_c$ a new minimizer, say at $\eta = \eta_\star > 0$, occurs, whence

$$a_2(T_c)\,\eta_\star^2 + a_3(T_c)\eta_\star^3 + a_4(T_c)\eta_\star^4 = 0.$$

The existence of a global minimum different from $\eta = 0$ below the critical temperature translates into

$$\min_{\eta\in\mathbb{R}} a_2(T)\,\eta^2 + a_3(T)\eta^3 + a_4(T)\eta^4 < 0 \qquad \text{for all } T < T_c.$$

An example of coefficients satisfying all these conditions is, for instance:

$T_c := 2, \qquad a_2(T) := 1,$

a_3 a smooth and increasing function such that $a_3(T) < 2$ for all $T \in \mathbb{R}$ with $a_3(T) = T - 4$ for all $T \leq 5$,

$a_4(T) := 1 \qquad \text{and} \qquad \eta_\star := 1.$

The first-order theory of phase transitions

Also, we assume that at $T = T_c$ a new minimizer, say at $\eta = \eta_\star > 0$, occurs, whence

$$a_2(T_c)\,\eta_\star^2 + a_3(T_c)\eta_\star^3 + a_4(T_c)\eta_\star^4 = 0.$$

The existence of a global minimum different from $\eta = 0$ below the critical temperature translates into

$$\min_{\eta\in\mathbb{R}} a_2(T)\,\eta^2 + a_3(T)\eta^3 + a_4(T)\eta^4 < 0 \qquad \text{for all } T < T_c.$$

An example of coefficients satisfying all these conditions is, for instance:

$T_c := 2, \qquad a_2(T) := 1,$

a_3 a smooth and increasing function such that $a_3(T) < 2$ for all $T \in \mathbb{R}$

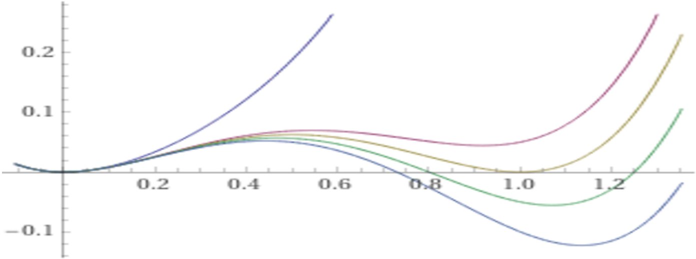
with $a_3(T) = T - 4$ for all $T \leq 5$,

$a_4(T) := 1 \qquad \text{and} \qquad \eta_\star := 1.$

The first-order theory of phase transitions

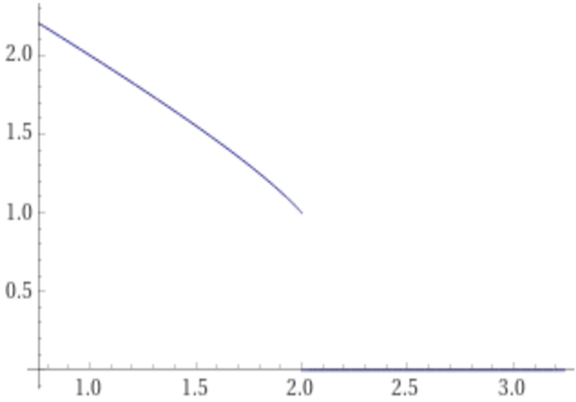
In this model case, the global minima of the free energy are described by

$$\begin{cases} \{0\} & \text{if } T > 2, \\ \{0, 1\} & \text{if } T = 2, \\ \left\{ \dfrac{12 - 3T + \sqrt{9T^2 - 72T + 112}}{8} \right\} & \text{if } T < 2, \end{cases}$$



The first-order theory of phase transitions

The bifurcation diagram for the minimizers is therefore



The first-order theory of phase transitions

This shows a discontinuous jump at the critical temperature for the minima of the free energy, which corresponds to the abrupt formation of a new stable phase.

The free energy as a function of temperature takes the form

$$\mathcal{E}(T) := \begin{cases} 0 & \text{if } T \geq T_c, \\ a_2(T)\,\eta^2(T) + a_3(T)\,\eta^3(T) + a_4(T)\,\eta^4(T) & \text{if } T < T_c. \end{cases}$$

As a result,

$$\begin{aligned} &\lim_{T \nearrow T_c} \mathcal{E}'(T) \\ &= a_2'(T_c)\,\eta^2(T_c) + a_3'(T_c)\,\eta^3(T_c) + a_4'(T_c)\,\eta^4(T_c) \\ &\qquad\qquad + \Big(2a_2(T_c)\,\eta(T_c) + 3a_3(T_c)\,\eta^2(T_c) + 4a_4(T_c)\,\eta^3(T_c)\Big)\eta'(T_c) \\ &= a_2'(T_c)\,\eta^2(T_c) + a_3'(T_c)\,\eta^3(T_c) + a_4'(T_c)\,\eta^4(T_c). \end{aligned}$$

The first-order theory of phase transitions

This shows a discontinuous jump at the critical temperature for the minima of the free energy, which corresponds to the abrupt formation of a new stable phase.

The free energy as a function of temperature takes the form

$$\mathcal{E}(T) := \begin{cases} 0 & \text{if } T \geq T_c, \\ a_2(T)\,\eta^2(T) + a_3(T)\,\eta^3(T) + a_4(T)\,\eta^4(T) & \text{if } T < T_c. \end{cases}$$

As a result,

$$\begin{aligned} &\lim_{T \nearrow T_c} \mathcal{E}'(T) \\ &= a_2'(T_c)\,\eta^2(T_c) + a_3'(T_c)\,\eta^3(T_c) + a_4'(T_c)\,\eta^4(T_c) \\ &\qquad + \Big(2a_2(T_c)\,\eta(T_c) + 3a_3(T_c)\,\eta^2(T_c) + 4a_4(T_c)\,\eta^3(T_c)\Big)\eta'(T_c) \\ &= a_2'(T_c)\,\eta^2(T_c) + a_3'(T_c)\,\eta^3(T_c) + a_4'(T_c)\,\eta^4(T_c). \end{aligned}$$

The first-order theory of phase transitions

This shows a discontinuous jump at the critical temperature for the minima of the free energy, which corresponds to the abrupt formation of a new stable phase.

The free energy as a function of temperature takes the form

$$\mathcal{E}(T) := \begin{cases} 0 & \text{if } T \geq T_c, \\ a_2(T)\,\eta^2(T) + a_3(T)\,\eta^3(T) + a_4(T)\,\eta^4(T) & \text{if } T < T_c. \end{cases}$$

As a result,

$$\begin{aligned} &\lim_{T \nearrow T_c} \mathcal{E}'(T) \\ &= a_2'(T_c)\,\eta^2(T_c) + a_3'(T_c)\,\eta^3(T_c) + a_4'(T_c)\,\eta^4(T_c) \\ &\qquad + \Big(2a_2(T_c)\,\eta(T_c) + 3a_3(T_c)\,\eta^2(T_c) + 4a_4(T_c)\,\eta^3(T_c)\Big)\eta'(T_c) \\ &= a_2'(T_c)\,\eta^2(T_c) + a_3'(T_c)\,\eta^3(T_c) + a_4'(T_c)\,\eta^4(T_c). \end{aligned}$$

The first-order theory of phase transitions

One can check that this quantity is always nonnegative (and strictly positive in “nondegenerate” cases).

The strict inequality corresponds to the typical situations in the so-called first-order phase transitions, in which the derivative of the free energy with respect to temperature is discontinuous at the critical temperature, which in turn corresponds to a physical situation in which a latent heat is emitted or absorbed by the system when the phase change occurs.

The first-order theory of phase transitions

One can check that this quantity is always nonnegative (and strictly positive in "nondegenerate" cases).

The strict inequality corresponds to the typical situations in the so-called first-order phase transitions, in which the derivative of the free energy with respect to temperature is discontinuous at the critical temperature, which in turn corresponds to a physical situation in which a latent heat is emitted or absorbed by the system when the phase change occurs.

The first-order theory of phase transitions

One can check that this quantity is always nonnegative (and strictly positive in “nondegenerate” cases).

The strict inequality corresponds to the typical situations in the so-called first-order phase transitions, in which the derivative of the free energy with respect to temperature is discontinuous at the critical temperature, which in turn corresponds to a physical situation in which a latent heat is emitted or absorbed by the system when the phase change occurs.

The first-order theory of phase transitions

In the model case,

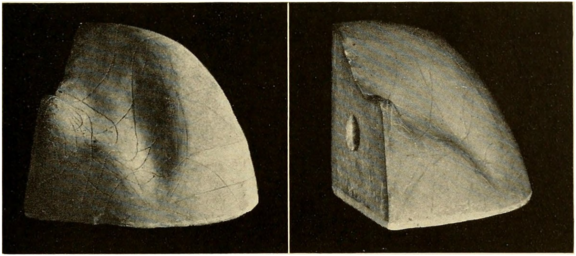
$$\mathcal{E}(T) = \begin{cases} 0 & \text{if } T \geq 2, \\ \frac{1}{512}\Big[\sqrt{9T^2 - 72T + 112}\big(9T^3 - 108T^2 + 400T - 448\big) \\ \quad -27T^4 + 432T^3 - 2448T^2 + 5760T - 4736\Big] & \text{if } T < 2 \end{cases}$$

and so

$$\lim_{T \nearrow 2} \mathcal{E}'(T) = 1 > 0 = \lim_{T \searrow 2} \mathcal{E}'(T),$$

showing the occurrence of the discontinuity at the critical temperature of the first derivative of the free energy with respect to temperature.

Putting our hands into phase transitions



Inspired by Gibbs' work, in 1874 Maxwell spent an entire winter to make a three-dimensional clay sculpture (also replicated in several plaster casts, one of which was sent by Maxwell to Gibbs as a gift) representing the energy of a fictitious substance with respect to volume and entropy.

Putting our hands into phase transitions

In this sculpture, one can recognize the principal features of phase transitions and latent heat formation via simple geometrical operations, such as placing a flat sheet of glass to mimic the tangent plane of the surface, or placing the model in sunlight and tracing the curve when the rays graze the surface.

In 2005, the United States Postal Service issued a 37 cents commemorative postage stamp honoring Gibbs. Next to Gibbs's portrait, the stamp features a diagram illustrating a thermodynamic surface. Also, an almost invisible microprinting on the collar of Gibbs' portrait depicts the equation $d\varepsilon = t\,d\eta - p\,dv$ (in Gibbs' notation, ε stands for internal energy, t for temperature, η for entropy, p for pressure and v for volume).

Putting our hands into phase transitions

In this sculpture, one can recognize the principal features of phase transitions and latent heat formation via simple geometrical operations, such as placing a flat sheet of glass to mimic the tangent plane of the surface, or placing the model in sunlight and tracing the curve when the rays graze the surface.

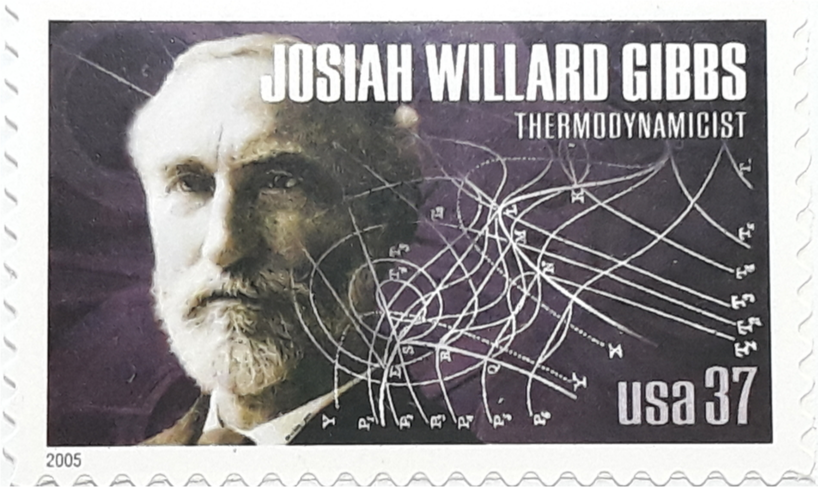
In 2005, the United States Postal Service issued a 37 cents commemorative postage stamp honoring Gibbs. Next to Gibbs's portrait, the stamp features a diagram illustrating a thermodynamic surface. Also, an almost invisible microprinting on the collar of Gibbs' portrait depicts the equation $d\varepsilon = t\,d\eta - p\,dv$ (in Gibbs' notation, ε stands for internal energy, t for temperature, η for entropy, p for pressure and v for volume).

Putting our hands into phase transitions

In this sculpture, one can recognize the principal features of phase transitions and latent heat formation via simple geometrical operations, such as placing a flat sheet of glass to mimic the tangent plane of the surface, or placing the model in sunlight and tracing the curve when the rays graze the surface.

In 2005, the United States Postal Service issued a 37 cents commemorative postage stamp honoring Gibbs. Next to Gibbs's portrait, the stamp features a diagram illustrating a thermodynamic surface. Also, an almost invisible microprinting on the collar of Gibbs' portrait depicts the equation $d\varepsilon = t\,d\eta - p\,dv$ (in Gibbs' notation, ε stands for internal energy, t for temperature, η for entropy, p for pressure and v for volume).

Putting our hands into phase transitions



Putting our hands into phase transitions

Interfacial energy of phase transitions

The description of phase models so far focused only on the favorable configurations of the free energy which support one phase over the other, but

- how are two different coexisting phases separated?
- What is the geometry is of the domains corresponding to each phase?

Interfacial energy of phase transitions

For this, note that the coexistence of two phases occurs when they both attain the minimal value of the free energy: this is precisely the case of first-order phase transitions at the critical temperature and of second-order phase transitions at the critical temperature or below it.

In principle, when two minima of the free energy occur at the same level, the two phases are equally favorable from an energetic point of view, hence any configuration in which any point of the state lies in any of the two phases is as good as any other.

This however is in contradiction with common experience, since in many phenomena the change of phase between different regions of the space occurs in very specific regions.

Interfacial energy of phase transitions

For this, note that the coexistence of two phases occurs when they both attain the minimal value of the free energy: this is precisely the case of first-order phase transitions at the critical temperature and of second-order phase transitions at the critical temperature or below it.

In principle, when two minima of the free energy occur at the same level, the two phases are equally favorable from an energetic point of view, hence any configuration in which any point of the state lies in any of the two phases is as good as any other.

This however is in contradiction with common experience, since in many phenomena the change of phase between different regions of the space occurs in very specific regions.

Interfacial energy of phase transitions

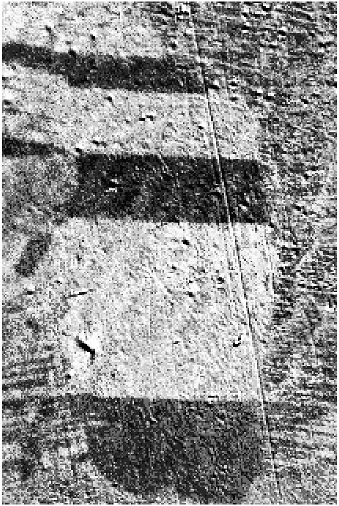
For this, note that the coexistence of two phases occurs when they both attain the minimal value of the free energy: this is precisely the case of first-order phase transitions at the critical temperature and of second-order phase transitions at the critical temperature or below it.

In principle, when two minima of the free energy occur at the same level, the two phases are equally favorable from an energetic point of view, hence any configuration in which any point of the state lies in any of the two phases is as good as any other.

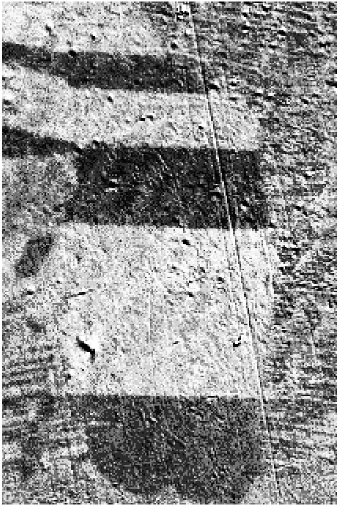
This however is in contradiction with common experience, since in many phenomena the change of phase between different regions of the space occurs in very specific regions.

Interfacial energy of phase transitions

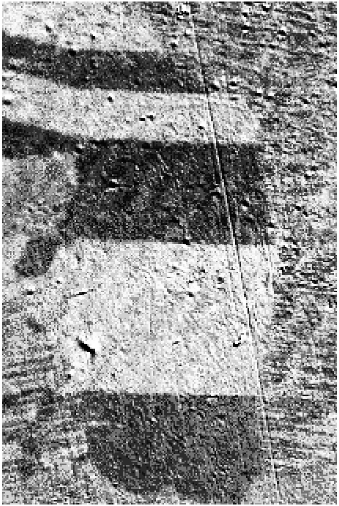
Interfacial energy of phase transitions



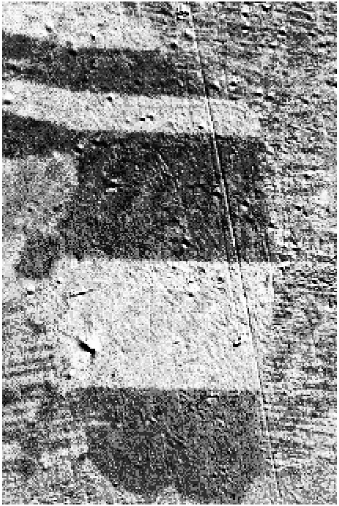
Interfacial energy of phase transitions



Interfacial energy of phase transitions

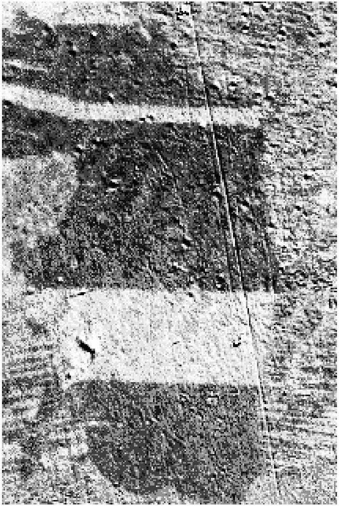


Interfacial energy of phase transitions

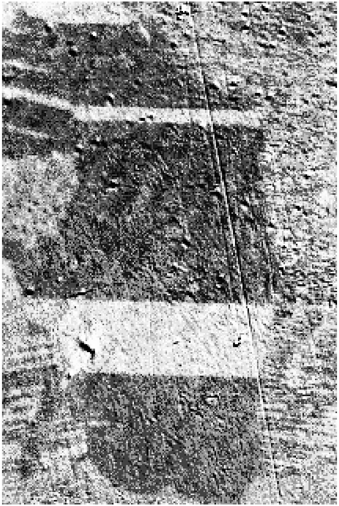


Interfacial energy of phase transitions

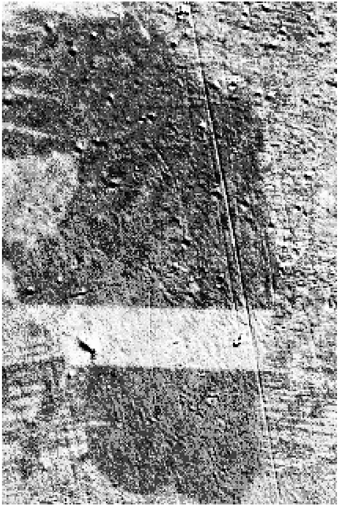
Interfacial energy of phase transitions



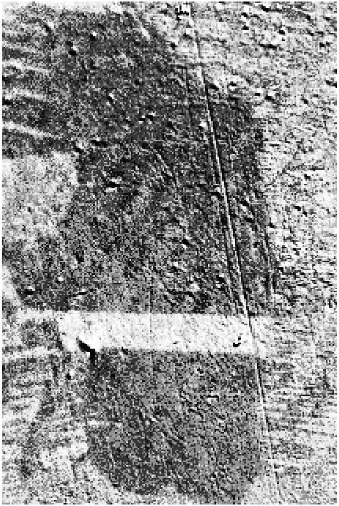
Interfacial energy of phase transitions



Interfacial energy of phase transitions



Interfacial energy of phase transitions



Interfacial energy of phase transitions

Interfacial energy of phase transitions

Interfacial energy of phase transitions

Interfacial energy of phase transitions

Interfacial energy of phase transitions

The model of “pure phases” is possibly “too abrupt”, since small fluctuations produce values of the state parameter which are not precisely equal to either of the two phases.

To model these fluctuations, one can consider the order parameter η as a function of the continuous spatial coordinates and assume that the fluctuations are the byproduct of the mutual interaction between regions of spaces corresponding to a different state parameter.

A crude, but perhaps efficient, model is to assume that the effective energy of the system comes from the free energy, plus an interaction term between different phases.

Interfacial energy of phase transitions

The model of “pure phases” is possibly “too abrupt”, since small fluctuations produce values of the state parameter which are not precisely equal to either of the two phases.

To model these fluctuations, one can consider the order parameter η as a function of the continuous spatial coordinates and assume that the fluctuations are the byproduct of the mutual interaction between regions of spaces corresponding to a different state parameter.

A crude, but perhaps efficient, model is to assume that the effective energy of the system comes from the free energy, plus an interaction term between different phases.

Interfacial energy of phase transitions

The model of “pure phases” is possibly “too abrupt”, since small fluctuations produce values of the state parameter which are not precisely equal to either of the two phases.

To model these fluctuations, one can consider the order parameter η as a function of the continuous spatial coordinates and assume that the fluctuations are the byproduct of the mutual interaction between regions of spaces corresponding to a different state parameter.

A crude, but perhaps efficient, model is to assume that the effective energy of the system comes from the free energy, plus an interaction term between different phases.

Interfacial energy of phase transitions

The introduction of this interaction term dates back at least to Van der Waals, who considered molecular interactions as averaged over long ranges. One of the additional benefits of such an interaction term is to penalize the unnecessary changes of phase and favor, among all the configurations which minimize the free energy, the ones which present a “minimal interface” between regions with different phases.

The precise notion of minimal interface certainly depends on the additional interaction term that one takes into account, thus we describe now some specific choices of interest. Considering the spatial domain the whole of $\mathbb{R}^n$, one can take into account an interaction energy of the form

$$\iint_{\mathbb{R}^n\times\mathbb{R}^n} (\eta(x)-\eta(y))^2\,\mathcal{K}(x,y)\,dx\,dy.$$

Interfacial energy of phase transitions

The introduction of this interaction term dates back at least to Van der Waals, who considered molecular interactions as averaged over long ranges. One of the additional benefits of such an interaction term is to penalize the unnecessary changes of phase and favor, among all the configurations which minimize the free energy, the ones which present a "minimal interface" between regions with different phases.

The precise notion of minimal interface certainly depends on the additional interaction term that one takes into account, thus we describe now some specific choices of interest. Considering the spatial domain the whole of $\mathbb{R}^n$, one can take into account an interaction energy of the form

$$\iint_{\mathbb{R}^n\times\mathbb{R}^n} (\eta(x)-\eta(y))^2\,\mathcal{K}(x,y)\,dx\,dy.$$

Interfacial energy of phase transitions

Ideally, one may want to determine a precise interaction kernel from general first principles.

However, due to the complexity of natural phenomena, in many concrete situations, the precise determination of an interaction kernel may be based on phenomenological considerations, interpretation of experimental data, or even, more frequently than one may think at a first, on the opportunity of finding useful computational simplifications.

As an example of “convenient choice” of an interaction kernel, we recall the fact that, in the development of his new theory of gases based on statistical physics, James Clerk Maxwell introduced an interaction kernel based on the inverse fifth power of the molecular distance. Actually, it seems that the model was possibly taking into account the general case of the κth power of the distance: according to several historical reconstructions, *Maxwell admitted that this choice of* $\kappa = 5$ *less from the physical consequences of the choice than from the attractiveness of the possibility of explicit integration.*

Interfacial energy of phase transitions

Ideally, one may want to determine a precise interaction kernel from general first principles.

However, due to the complexity of natural phenomena, in many concrete situations, the precise determination of an interaction kernel may be based on phenomenological considerations, interpretation of experimental data, or even, more frequently than one may think at a first, on the opportunity of finding useful computational simplifications.

As an example of "convenient choice" of an interaction kernel, we recall the fact that, in the development of his new theory of gases based on statistical physics, James Clerk Maxwell introduced an interaction kernel based on the inverse fifth power of the molecular distance. Actually, it seems that the model was possibly taking into account the general case of the κth power of the distance: according to several historical reconstructions, *Maxwell admitted that this choice of* $\kappa = 5$ *less from the physical consequences of the choice than from the attractiveness of the possibility of explicit integration.*

Interfacial energy of phase transitions

Ideally, one may want to determine a precise interaction kernel from general first principles.

However, due to the complexity of natural phenomena, in many concrete situations, the precise determination of an interaction kernel may be based on phenomenological considerations, interpretation of experimental data, or even, more frequently than one may think at a first, on the opportunity of finding useful computational simplifications.

As an example of “convenient choice” of an interaction kernel, we recall the fact that, in the development of his new theory of gases based on statistical physics, James Clerk Maxwell introduced an interaction kernel based on the inverse fifth power of the molecular distance. Actually, it seems that the model was possibly taking into account the general case of the κth power of the distance: according to several historical reconstructions, *Maxwell admitted that this choice of $\kappa = 5$ less from the physical consequences of the choice than from the attractiveness of the possibility of explicit integration.*

Interfacial energy of phase transitions

If we assume the kernel $\mathcal{K}$ to be symmetric under translations and rotations, we have that

$$\begin{aligned}\mathcal{K}(x,y) &= \mathcal{K}(x-y,y-y)\\ &= \mathcal{K}(x-y,0) = \mathcal{K}(|x-y|e_1,0) =: K(|x-y|).\end{aligned}$$

Interfacial energy of phase transitions

If additionally the kernel is short-range, i.e. it vanishes whenever $|x-y| \geq \varrho$, for some small $\varrho > 0$,

$$\begin{aligned}
&\int_{\mathbb{R}^n} \left[\int_{B_\varrho(x)} (\eta(x) - \eta(y))^2 \, K(|x-y|) \, dy \right] dx \\
&= \int_{\mathbb{R}^n} \left[\int_{B_\varrho(x)} \Big(\nabla\eta(x) \cdot (x-y) + O(\sigma(x) \, |x-y|^2) \Big)^2 \, K(|x-y|) \, dy \right] dx \\
&= \int_{\mathbb{R}^n} \left[\int_{B_\varrho} \Big(\nabla\eta(x) \cdot z + O\big(\sigma(x) \, |z|^2\big) \Big)^2 \, K(|z|) \, dz \right] dx \\
&= \int_{\mathbb{R}^n} \left[\int_{B_\varrho} \big(\nabla\eta(x) \cdot z \big)^2 \, K(|z|) \, dz \right] dx + \int_{\mathbb{R}^n} \left[\int_{B_\varrho} O(\sigma^2(x) \, |z|^3) \, K(|z|) \, dz \right] dx \\
&= \int_{\mathbb{R}^n} \left[\int_{B_\varrho} |\nabla\eta(x)|^2 \, z_1^2 \, K(|z|) \, dz \right] dx + \int_{\mathbb{R}^n} \left[\int_{B_\varrho} O(\sigma^2(x) \, |z|^3) \, K(|z|) \, dz \right] dx \\
&= C \int_{\mathbb{R}^n} |\nabla\eta(x)|^2 \, dx + O(C\varrho),
\end{aligned}$$

where $\sigma(x) := \|\eta\|_{C^2(B_\varrho(x))}$ and $C := \int_{\mathbb{R}^n} z_1^2 \, K(|z|) \, dz$.

Interfacial energy of phase transitions

Accordingly, for ϱ sufficiently small, the interaction term can be approximated by

$$\int_{\mathbb{R}^n} |\nabla\eta(x)|^2\,dx.$$

The coexistence of two phases for first-order phase transitions at the critical temperature and of second-order phase transitions at the critical temperature or below it produces, in a container Ω, the energy functional

$$\int_\Omega |\nabla\eta(x)|^2\,dx + \int_\Omega W(\eta(x))\,dx,$$

where W is a double-well potential (e.g., with η prescribed outside Ω, or equivalently along $\partial\Omega$).

Interfacial energy of phase transitions

Accordingly, for ϱ sufficiently small, the interaction term can be approximated by

$$\int_{\mathbb{R}^n} |\nabla \eta(x)|^2 \, dx.$$

The coexistence of two phases for first-order phase transitions at the critical temperature and of second-order phase transitions at the critical temperature or below it produces, in a container Ω, the energy functional

$$\int_{\Omega} |\nabla \eta(x)|^2 \, dx + \int_{\Omega} W(\eta(x)) \, dx,$$

where W is a double-well potential (e.g., with η prescribed outside Ω, or equivalently along $\partial\Omega$).

Interfacial energy of phase transitions

The phase separation in this case is dictated by the usual surface tension aiming at making the interface a codimension 1 surface with the least possible $(n-1)$-dimensional area.

To see this, at least heuristically, one considers a rescaling of the energy in which the gradient term is explicitly a penalization of the double-well potential responsible of the phase separation.

Interfacial energy of phase transitions

The phase separation in this case is dictated by the usual surface tension aiming at making the interface a codimension 1 surface with the least possible $(n-1)$-dimensional area.

To see this, at least heuristically, one considers a rescaling of the energy in which the gradient term is explicitly a penalization of the double-well potential responsible of the phase separation.

Interfacial energy of phase transitions

That is, for a small parameter $\varepsilon > 0$, one takes into account the energy functional

$$\varepsilon \int_\Omega |\nabla \eta(x)|^2 \, dx + \frac{1}{\varepsilon} \int_\Omega W(\eta(x)) \, dx.$$

By the Cauchy-Schwarz Inequality and the Coarea Formula, one can bound this quantity from below by

$$\begin{aligned} \int_\Omega |\nabla \eta(x)| \sqrt{W(\eta(x))} \, dx &= \int_{-1}^{1} \left[\int_{\Omega \cap \{\eta(x)=\tau\}} \sqrt{W(\tau)} \, d\mathcal{H}_x^{n-1} \right] d\tau \\ &= \int_{-1}^{1} \sqrt{W(\tau)} \, \mathcal{H}^{n-1}\Big(\Omega \cap \{\eta = \tau\}\Big) \, d\tau, \end{aligned}$$

where $\mathcal{H}^{n-1}$ denotes the $(n-1)$-dimensional Hausdorff measure.

Interfacial energy of phase transitions

Also, for small ε, we may think that the energy minimizers try to "optimize" the above lower bound and to sit in the zeros (or close to the zeros) of the double-well potential as much as possible.

Therefore, for small ε, the minimal energy is expected to be related to

$$c\,\mathcal{H}^{n-1}\Big(\Omega\cap(\partial E)\Big) \qquad \text{where} \qquad c := \int_{-1}^{1}\sqrt{W(\tau)}\,d\tau,$$

being E a set in which the order parameter is "essentially" equal to $+1$ and the complement of E a set in which the order parameter is "essentially" equal to -1.

Interfacial energy of phase transitions

Also, for small ε, we may think that the energy minimizers try to "optimize" the above lower bound and to sit in the zeros (or close to the zeros) of the double-well potential as much as possible.

Therefore, for small ε, the minimal energy is expected to be related to

$$c\,\mathcal{H}^{n-1}\big(\Omega \cap (\partial E)\big) \qquad \text{where} \qquad c := \int_{-1}^{1} \sqrt{W(\tau)}\,d\tau,$$

being E a set in which the order parameter is "essentially" equal to $+1$ and the complement of E a set in which the order parameter is "essentially" equal to -1.

Interfacial energy of phase transitions

For long-range interactions, the gradient approximation is not available anymore and we have a nonlocal energy term of the form

$$\iint_{\mathcal{Q}(\Omega)} (\eta(x)-\eta(y))^2\, K(|x-y|)\,dx\,dy + \int_\Omega W(\eta(x))\,dx,$$

where

$$\mathcal{Q}(\Omega) := (\Omega\times\Omega)\cup(\Omega^c\times\Omega)\cup(\Omega\times\Omega^c),$$

being $\Omega^c := \mathbb{R}^n \setminus \Omega$.

The notation $\mathcal{Q}(\Omega)$ stands for a "cross-shaped set" ("$\mathcal{Q}$" stands for cross, since "C" is used for constants and "K" for kernels!).

Interfacial energy of phase transitions

For long-range interactions, the gradient approximation is not available anymore and we have a nonlocal energy term of the form

$$\iint_{\mathcal{Q}(\Omega)} (\eta(x) - \eta(y))^2 \, K(|x-y|) \, dx \, dy + \int_\Omega W(\eta(x)) \, dx,$$

where

$$\mathcal{Q}(\Omega) := (\Omega \times \Omega) \cup (\Omega^c \times \Omega) \cup (\Omega \times \Omega^c),$$

being $\Omega^c := \mathbb{R}^n \setminus \Omega$.

The notation $\mathcal{Q}(\Omega)$ stands for a "cross-shaped set" ("$\mathcal{Q}$" stands for cross, since "C" is used for constants and "K" for kernels!).

Interfacial energy of phase transitions

The intuition behind $Q(\Omega)$ is that we are prescribing here the order parameter η outside the domain Ω, which is the "global" counterpart of prescribing η along $\partial\Omega$.

Accordingly, the long-range energy functional should account for all the configurations which involve the values of the state parameter in Ω: whatever piece of energy containing only the values of the state parameter outside Ω is prescribed, whence does not influence energy minimization (interestingly, in this way, one considers the energy confined outside the domain as "constant", even if this constant can actually be infinite!).

In this spirit, the cross-shaped set accounts for all the phase interactions in which at least one of the sites is located in Ω.

Interfacial energy of phase transitions

The intuition behind $Q(\Omega)$ is that we are prescribing here the order parameter η outside the domain Ω, which is the “global” counterpart of prescribing η along $\partial\Omega$.

Accordingly, the long-range energy functional should account for all the configurations which involve the values of the state parameter in Ω: whatever piece of energy containing only the values of the state parameter outside Ω is prescribed, whence does not influence energy minimization (interestingly, in this way, one considers the energy confined outside the domain as “constant”, even if this constant can actually be infinite!).

In this spirit, the cross-shaped set accounts for all the phase interactions in which at least one of the sites is located in Ω.

Interfacial energy of phase transitions

The intuition behind $Q(\Omega)$ is that we are prescribing here the order parameter η outside the domain Ω, which is the "global" counterpart of prescribing η along $\partial\Omega$.

Accordingly, the long-range energy functional should account for all the configurations which involve the values of the state parameter in Ω: whatever piece of energy containing only the values of the state parameter outside Ω is prescribed, whence does not influence energy minimization (interestingly, in this way, one considers the energy confined outside the domain as "constant", even if this constant can actually be infinite!).

In this spirit, the cross-shaped set accounts for all the phase interactions in which at least one of the sites is located in Ω.

Interfacial energy of phase transitions

Special cases of kernels K are the ones which are positively homogeneous of some degree d, that is $K(|tz|) = t^d K(|z|)$ for all $z \in \mathbb{R}^n \setminus \{0\}$ and $t \in (0, +\infty)$.

Note that the degree d cannot be arbitrarily chosen in the reals, since to make sense of the interaction energy it is desirable to have it finite at least when $\eta \in C_0^\infty(\Omega, [0,2])$ with $\eta(x) = 2 - |x - x_0|^2$ for all $x \in B_r(x_0)$, for some small $r \in (0,1)$ such that $B_{2r}(x_0) \subseteq \Omega \subseteq B_{1/r}(x_0)$. Hence,

$$d \in (-n-2, -n).$$

Interfacial energy of phase transitions

Special cases of kernels K are the ones which are positively homogeneous of some degree d, that is $K(|tz|) = t^d K(|z|)$ for all $z \in \mathbb{R}^n \setminus \{0\}$ and $t \in (0, +\infty)$.

Note that the degree d cannot be arbitrarily chosen in the reals, since to make sense of the interaction energy it is desirable to have it finite at least when $\eta \in C_0^\infty(\Omega, [0, 2])$ with $\eta(x) = 2 - |x - x_0|^2$ for all $x \in B_r(x_0)$, for some small $r \in (0, 1)$ such that $B_{2r}(x_0) \subseteq \Omega \subseteq B_{1/r}(x_0)$. Hence,

$$d \in (-n - 2, -n).$$

Interfacial energy of phase transitions

So, necessarily

$$K(|z|) = |z|^d \, K(|e_1|) = \frac{1}{|z|^{n+\alpha}},$$

for some $\alpha \in (0, 2)$ and

$$\iint_{Q(\Omega)} \frac{(\eta(x) - \eta(y))^2}{|x - y|^{n+\alpha}} \, dx \, dy + \int_\Omega W(\eta(x)) \, dx.$$

Interfacial energy of phase transitions

When $\alpha \in (0,1)$, the minimizers can be easily related to a geometric minimization problem, since if

$$\eta(x) = \chi_E(x) - \chi_{E^c}(x) = \begin{cases} 1 & \text{if } x \in E, \\ -1 & \text{if } x \in E^c, \end{cases}$$

for some $E \subseteq \mathbb{R}^n$, then the energy functional boils down to

$$\iint_{Q(\Omega)\cap(E\times E^c)} \frac{4}{|x-y|^{n+\alpha}}\,dx\,dy = \operatorname{Per}_\alpha(E,\Omega).$$

Interestingly, when $\alpha \in [1,2)$, the limit interfaces will be instead related to the minimizers of the classical perimeter, showing a remarkable "localization effect for nonlocal energies" when the interaction parameter α is larger than or equal to 1.

Interfacial energy of phase transitions

When $\alpha \in (0, 1)$, the minimizers can be easily related to a geometric minimization problem, since if

$$\eta(x) = \chi_E(x) - \chi_{E^c}(x) = \begin{cases} 1 & \text{if } x \in E, \\ -1 & \text{if } x \in E^c, \end{cases}$$

for some $E \subseteq \mathbb{R}^n$, then the energy functional boils down to

$$\iint_{Q(\Omega)\cap(E\times E^c)} \frac{4}{|x-y|^{n+\alpha}}\, dx\, dy = \mathrm{Per}_\alpha(E, \Omega).$$

Interestingly, when $\alpha \in [1, 2)$, the limit interfaces will be instead related to the minimizers of the classical perimeter, showing a remarkable "localization effect for nonlocal energies" when the interaction parameter α is larger than or equal to 1.

Interfacial energy of phase transitions

Road plan:

- Understand better the minimizers (and the critical points, and the quasi-minimizers, etc.) of the short-range and long-range phase transition energy functionals,
- Understand better the link between the short-range and long-range phase transition energy functionals with the limit geometric problem,
- Understand better the structure of the minimizers (and the critical points, and the quasi-minimizers, etc.) limit geometric problem,
- Understand better the influence of the structure of the minimizers (and the critical points, and the quasi-minimizers, etc.) limit geometric problem on the corresponding objects for the short-range and long-range phase transition energy functionals.

Interfacial energy of phase transitions

Road plan:

- Understand better the minimizers (and the critical points, and the quasi-minimizers, etc.) of the short-range and long-range phase transition energy functionals,
- Understand better the link between the short-range and long-range phase transition energy functionals with the limit geometric problem,
- Understand better the structure of the minimizers (and the critical points, and the quasi-minimizers, etc.) limit geometric problem,
- Understand better the influence of the structure of the minimizers (and the critical points, and the quasi-minimizers, etc.) limit geometric problem on the corresponding objects for the short-range and long-range phase transition energy functionals.

Interfacial energy of phase transitions

Road plan:

- Understand better the minimizers (and the critical points, and the quasi-minimizers, etc.) of the short-range and long-range phase transition energy functionals,
- Understand better the link between the short-range and long-range phase transition energy functionals with the limit geometric problem,
- Understand better the structure of the minimizers (and the critical points, and the quasi-minimizers, etc.) limit geometric problem,
- Understand better the influence of the structure of the minimizers (and the critical points, and the quasi-minimizers, etc.) limit geometric problem on the corresponding objects for the short-range and long-range phase transition energy functionals.

Interfacial energy of phase transitions

Road plan:

- Understand better the minimizers (and the critical points, and the quasi-minimizers, etc.) of the short-range and long-range phase transition energy functionals,
- Understand better the link between the short-range and long-range phase transition energy functionals with the limit geometric problem,
- Understand better the structure of the minimizers (and the critical points, and the quasi-minimizers, etc.) limit geometric problem,
- Understand better the influence of the structure of the minimizers (and the critical points, and the quasi-minimizers, etc.) limit geometric problem on the corresponding objects for the short-range and long-range phase transition energy functionals.

Interfacial energy of phase transitions

Road plan:

- Understand better the minimizers (and the critical points, and the quasi-minimizers, etc.) of the short-range and long-range phase transition energy functionals,
- Understand better the link between the short-range and long-range phase transition energy functionals with the limit geometric problem,
- Understand better the structure of the minimizers (and the critical points, and the quasi-minimizers, etc.) limit geometric problem,
- Understand better the influence of the structure of the minimizers (and the critical points, and the quasi-minimizers, etc.) limit geometric problem on the corresponding objects for the short-range and long-range phase transition energy functionals.

Minimal surfaces

The first and second variations of the perimeter functional can be explicitly computed in terms of the mean curvature H and of the norm c of the second fundamental form.

More precisely, one can consider a domain $\Omega \subset \mathbb{R}^n$, a set $E \subset \mathbb{R}^n$ and a function $\phi \in C_0^\infty(\Omega)$ such that ∂E is a hypersurface of class C^2 in the support S of ϕ.

Thus, we take the exterior normal vector ν to E in S, consider the vector field $\phi\nu$ (extended to 0 outside S).

Minimal surfaces

The first and second variations of the perimeter functional can be explicitly computed in terms of the mean curvature H and of the norm c of the second fundamental form.

More precisely, one can consider a domain $\Omega \subset \mathbb{R}^n$, a set $E \subset \mathbb{R}^n$ and a function $\phi \in C_0^\infty(\Omega)$ such that ∂E is a hypersurface of class C^2 in the support S of ϕ.

Thus, we take the exterior normal vector ν to E in S, consider the vector field $\phi\nu$ (extended to 0 outside S).

Minimal surfaces

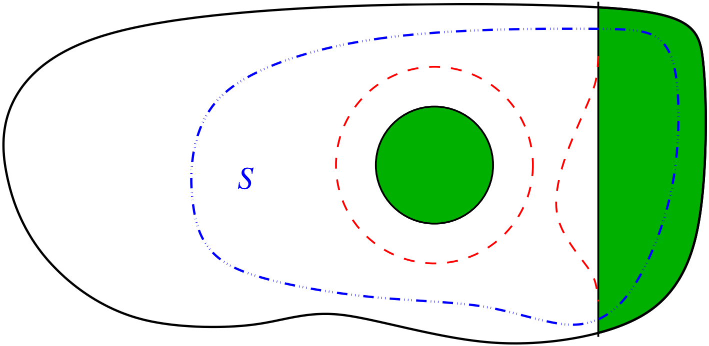
The first and second variations of the perimeter functional can be explicitly computed in terms of the mean curvature H and of the norm c of the second fundamental form.

More precisely, one can consider a domain $\Omega \subset \mathbb{R}^n$, a set $E \subset \mathbb{R}^n$ and a function $\phi \in C_0^\infty(\Omega)$ such that ∂E is a hypersurface of class C^2 in the support S of ϕ.

Thus, we take the exterior normal vector ν to E in S, consider the vector field $\phi\nu$ (extended to 0 outside S).

Minimal surfaces

We denote by E_t the flow of the set E along this vector field.



Minimal surfaces

Then, as $t \to 0$,

$$\mathrm{Per}(E_t, \Omega) = \mathrm{Per}(E, \Omega) + t \int_{\partial E} H(x)\, \phi(x)\, d\mathcal{H}^{n-1}_x$$
$$+ \frac{t^2}{2} \int_{\partial E} \Big(|\nabla_T \phi(x)|^2 - (c^2(x) - H^2(x))\phi^2(x) \Big)\, d\mathcal{H}^{n-1}_x + o(t),$$

where the tangential gradient is given by

$$\nabla_T \phi = \nabla\phi - (\nabla\phi \cdot \nu)\nu.$$

Minimal surfaces

Therefore, a critical point for the perimeter functional is (regularity allowing) a hypersurface with vanishing mean curvature and a minimizer satisfies additionally that

$$\int_{\partial E} \left(|\nabla_T \phi(x)|^2 - c^2(x)\, \phi^2(x) \right) d\mathcal{H}_x^{n-1} \geqslant 0$$

for every test function $\phi \in C_0^\infty(\Omega)$.

One says that a vanishing mean curvature hypersurface is stable if the latter condition is satisfied (in particular, local minimizers are stable).

Minimal surfaces

Therefore, a critical point for the perimeter functional is (regularity allowing) a hypersurface with vanishing mean curvature and a minimizer satisfies additionally that

$$\int_{\partial E} \Big(|\nabla_T \phi(x)|^2 - c^2(x)\,\phi^2(x) \Big)\, d\mathcal{H}^{n-1}_x \geqslant 0$$

for every test function $\phi \in C_0^\infty(\Omega)$.

One says that a vanishing mean curvature hypersurface is stable if the latter condition is satisfied (in particular, local minimizers are stable).

Minimal surfaces

One of the chief results in the classical theory of minimal surfaces is that:

Theorem (Simons 1968 (version 1))

Perimeter minimizers are smooth in dimension $n \leq 7$.

The dimensional assumption is optimal, since minimal cones occur in dimension $n \geq 8$, as constructed by [Bombieri, De Giorgi, Giusti 1969].

Minimal surfaces

One of the chief results in the classical theory of minimal surfaces is that:

Theorem (Simons 1968 (version 1))

Perimeter minimizers are smooth in dimension $n \leq 7$.

The dimensional assumption is optimal, since minimal cones occur in dimension $n \geq 8$, as constructed by [Bombieri, De Giorgi, Giusti 1969].

Minimal surfaces

The result by Simons is proved by reducing, after a blow-up procedure and a dimensional reduction, to the case in which the minimal surface E under consideration is a cone (namely, if $p \in E$ then $tp \in E$ for all $t > 0$) and its only possible singularity is at the origin.

Thus, in this setting, the result by Simons (version 1) is a consequence of

Theorem (Simons 1968 (version 2))

A vanishing mean curvature cone cone in $\mathbb{R}^n$ with $n \leq 7$ which is stable and smooth outside the origin is necessarily a halfplane.

Minimal surfaces

The result by Simons is proved by reducing, after a blow-up procedure and a dimensional reduction, to the case in which the minimal surface E under consideration is a cone (namely, if $p \in E$ then $tp \in E$ for all $t > 0$) and its only possible singularity is at the origin.

Thus, in this setting, the result by Simons (version 1) is a consequence of

Theorem (Simons 1968 (version 2))

A vanishing mean curvature cone cone in $\mathbb{R}^n$ with $n \leq 7$ which is stable and smooth outside the origin is necessarily a halfplane.

Minimal surfaces

The proof of this relies on a beautiful inequality of geometric type:

Theorem (Simons 1968 (version 3))

At every regular point of a cones E with zero mean curvature (not necessarily stable), it holds that

$$\frac{\Delta_T c^2}{2} \geq |\nabla_T c|^2 + \frac{2c^2}{|x|^2} - c^4.$$

Here Δ_T is the Laplace-Beltrami operator, which can be defined, for instance, in the distributional sense via the tangential gradient, for smooth and compactly supported functions f and g, by

$$\int_{\partial E} \Delta_T f(x)\, g(x)\, d\mathcal{H}_x^{n-1} = -\int_{\partial E} \nabla_T f(x) \cdot \nabla_T g(x)\, d\mathcal{H}_x^{n-1}.$$

Minimal surfaces

The proof of this relies on a beautiful inequality of geometric type:

Theorem (Simons 1968 (version 3))

At every regular point of a cones E with zero mean curvature (not necessarily stable), it holds that

$$\frac{\Delta_T c^2}{2} \geq |\nabla_T c|^2 + \frac{2c^2}{|x|^2} - c^4.$$

Here Δ_T is the Laplace-Beltrami operator, which can be defined, for instance, in the distributional sense via the tangential gradient, for smooth and compactly supported functions f and g, by

$$\int_{\partial E} \Delta_T f(x)\, g(x)\, d\mathcal{H}^{n-1}_x = -\int_{\partial E} \nabla_T f(x) \cdot \nabla_T g(x)\, d\mathcal{H}^{n-1}_x.$$

Minimal surfaces

Proof that version 3 implies version 2. We consider a test function ζ and exploit the stability inequality with $\phi := c\zeta$, finding that

$$\begin{aligned} 0 &\le \int_{\partial E} \Big(|\zeta(x)\nabla_T c(x) + c(x)\nabla_T\zeta(x)|^2 - c^4(x)\,\zeta^2(x) \Big)\, d\mathcal{H}^{n-1}_x \\ &= \int_{\partial E} \Big(|\zeta(x)\nabla_T c(x)|^2 + |c(x)\nabla_T\zeta(x)|^2 + \frac{1}{2}\nabla_T c^2(x)\cdot\nabla_T\zeta^2(x) - c^4(x)\,\zeta^2(x) \Big)\, d\mathcal{H}^{n-1}_x \\ &= \int_{\partial E} \Big(\zeta^2(x)|\nabla_T c(x)|^2 + c^2(x)|\nabla_T\zeta(x)|^2 - \frac{\Delta_T c^2(x)}{2}\zeta^2(x) - c^4(x)\,\zeta^2(x) \Big)\, d\mathcal{H}^{n-1}_x . \end{aligned}$$

Combining this with version 3, we infer that

$$\int_{\partial E} \frac{2c^2(x)\,\zeta^2(x)}{|x|^2}\, d\mathcal{H}^{n-1}_x \le \int_{\partial E} c^2(x)\,|\nabla_T\zeta(x)|^2\, d\mathcal{H}^{n-1}_x .$$

Minimal surfaces

Proof that version 3 implies version 2. We consider a test function ζ and exploit the stability inequality with $\phi := c\zeta$, finding that

$$\begin{aligned}
0 &\leq \int_{\partial E} \Big(|\zeta(x)\nabla_T c(x) + c(x)\nabla_T \zeta(x)|^2 - c^4(x)\,\zeta^2(x) \Big)\, d\mathcal{H}^{n-1}_x \\
&= \int_{\partial E} \left(|\zeta(x)\nabla_T c(x)|^2 + |c(x)\nabla_T \zeta(x)|^2 + \frac{1}{2}\nabla_T c^2(x)\cdot\nabla_T \zeta^2(x) - c^4(x)\,\zeta^2(x) \right) d\mathcal{H}^{n-1}_x \\
&= \int_{\partial E} \left(\zeta^2(x)|\nabla_T c(x)|^2 + c^2(x)|\nabla_T \zeta(x)|^2 - \frac{\Delta_T c^2(x)}{2}\zeta^2(x) - c^4(x)\,\zeta^2(x) \right) d\mathcal{H}^{n-1}_x .
\end{aligned}$$

Combining this with version 3, we infer that

$$\int_{\partial E} \frac{2c^2(x)\,\zeta^2(x)}{|x|^2}\, d\mathcal{H}^{n-1}_x \leq \int_{\partial E} c^2(x)\,|\nabla_T \zeta(x)|^2\, d\mathcal{H}^{n-1}_x .$$

Minimal surfaces

Proof that version 3 implies version 2. We consider a test function ζ and exploit the stability inequality with $\phi := c\zeta$, finding that

$$\begin{aligned}
0 &\leq \int_{\partial E} \Big(|\zeta(x)\nabla_T c(x) + c(x)\nabla_T \zeta(x)|^2 - c^4(x)\,\zeta^2(x) \Big)\, d\mathcal{H}^{n-1}_x \\
&= \int_{\partial E} \left(|\zeta(x)\nabla_T c(x)|^2 + |c(x)\nabla_T \zeta(x)|^2 + \frac{1}{2}\nabla_T c^2(x) \cdot \nabla_T \zeta^2(x) - c^4(x)\,\zeta^2(x) \right) d\mathcal{H}^{n-1}_x \\
&= \int_{\partial E} \left(\zeta^2(x)|\nabla_T c(x)|^2 + c^2(x)|\nabla_T \zeta(x)|^2 - \frac{\Delta_T c^2(x)}{2}\zeta^2(x) - c^4(x)\,\zeta^2(x) \right) d\mathcal{H}^{n-1}_x .
\end{aligned}$$

Combining this with version 3, we infer that

$$\int_{\partial E} \frac{2c^2(x)\,\zeta^2(x)}{|x|^2}\, d\mathcal{H}^{n-1}_x \leq \int_{\partial E} c^2(x)\,|\nabla_T \zeta(x)|^2\, d\mathcal{H}^{n-1}_x .$$

Minimal surfaces

Given $\alpha, \beta \in \mathbb{R}$, and $\varepsilon \in (0,1)$, we consider $\tau_\varepsilon \in C_0^\infty\Big(B_{2/\varepsilon} \setminus B_\varepsilon, [0,1]\Big)$ with $\tau_\varepsilon = 1$ in $B_{1/\varepsilon} \setminus B_{2\varepsilon}$ with

$$|\nabla \tau_\varepsilon| \leq \frac{4}{\varepsilon}\chi_{B_{2\varepsilon}\setminus B_\varepsilon} + 4\varepsilon\chi_{B_{2/\varepsilon}\setminus B_{1/\varepsilon}}.$$

Minimal surfaces

Let also $\zeta_\varepsilon := \tau_\varepsilon \varphi_\varepsilon$, with

$$\varphi_\varepsilon(x) := \frac{|x|^\alpha}{2}\left(\sqrt{\left(|x|^\beta - 1\right)^2 + \varepsilon} + |x|^\beta + 1\right).$$

The idea is to use ζ_ε as a test function and pass to the limit as $\varepsilon \searrow 0$. For this approximation method to work, we will need to choose appropriately the parameters α and β, which, in turn, will be possible only under the dimensional restriction.

For this, we note that

$$\lim_{\varepsilon \searrow 0} \zeta_\varepsilon(x) = \lim_{\varepsilon \searrow 0} \varphi_\varepsilon(x) = \frac{|x|^\alpha}{2}\left(\left||x|^\beta - 1\right| + |x|^\beta + 1\right) = \begin{cases} |x|^{\alpha+\beta} & \text{if } x \in \mathbb{R}^n \setminus B_1, \\ |x|^\alpha & \text{if } x \in B_1. \end{cases}$$

Minimal surfaces

Let also $\zeta_\varepsilon := \tau_\varepsilon \varphi_\varepsilon$, with

$$\varphi_\varepsilon(x) := \frac{|x|^\alpha}{2}\left(\sqrt{\left(|x|^\beta - 1\right)^2 + \varepsilon} + |x|^\beta + 1\right).$$

The idea is to use ζ_ε as a test function and pass to the limit as $\varepsilon \searrow 0$. For this approximation method to work, we will need to choose appropriately the parameters α and β, which, in turn, will be possible only under the dimensional restriction.

For this, we note that

$$\lim_{\varepsilon \searrow 0} \zeta_\varepsilon(x) = \lim_{\varepsilon \searrow 0} \varphi_\varepsilon(x) = \frac{|x|^\alpha}{2}\left(\left||x|^\beta - 1\right| + |x|^\beta + 1\right) = \begin{cases} |x|^{\alpha+\beta} & \text{if } x \in \mathbb{R}^n \setminus B_1, \\ |x|^\alpha & \text{if } x \in B_1. \end{cases}$$

Minimal surfaces

Let also $\zeta_\varepsilon := \tau_\varepsilon \varphi_\varepsilon$, with

$$\varphi_\varepsilon(x) := \frac{|x|^\alpha}{2}\left(\sqrt{\left(|x|^\beta - 1\right)^2 + \varepsilon} + |x|^\beta + 1\right).$$

The idea is to use ζ_ε as a test function and pass to the limit as $\varepsilon \searrow 0$. For this approximation method to work, we will need to choose appropriately the parameters α and β, which, in turn, will be possible only under the dimensional restriction.

For this, we note that

$$\lim_{\varepsilon \searrow 0} \zeta_\varepsilon(x) = \lim_{\varepsilon \searrow 0} \varphi_\varepsilon(x) = \frac{|x|^\alpha}{2}\left(\left||x|^\beta - 1\right| + |x|^\beta + 1\right) = \begin{cases} |x|^{\alpha+\beta} & \text{if } x \in \mathbb{R}^n \setminus B_1, \\ |x|^\alpha & \text{if } x \in B_1. \end{cases}$$

Minimal surfaces

Let also $\zeta_\varepsilon := \tau_\varepsilon \varphi_\varepsilon$, with

$$\varphi_\varepsilon(x) := \frac{|x|^\alpha}{2}\left(\sqrt{\left(|x|^\beta - 1\right)^2 + \varepsilon} + |x|^\beta + 1\right).$$

The idea is to use ζ_ε as a test function and pass to the limit as $\varepsilon \searrow 0$. For this approximation method to work, we will need to choose appropriately the parameters α and β, which, in turn, will be possible only under the dimensional restriction.

For this, we note that

$$\lim_{\varepsilon \searrow 0} \zeta_\varepsilon(x) = \lim_{\varepsilon \searrow 0} \varphi_\varepsilon(x) = \frac{|x|^\alpha}{2}\left(\left||x|^\beta - 1\right| + |x|^\beta + 1\right) = \begin{cases} |x|^{\alpha+\beta} & \text{if } x \in \mathbb{R}^n \setminus B_1, \\ |x|^\alpha & \text{if } x \in B_1. \end{cases}$$

Minimal surfaces

Thus, Fatou's Lemma entails that

$$\begin{aligned}\liminf_{\varepsilon\searrow 0}\int_{\partial E}\frac{2c^2(x)\,\zeta_\varepsilon^2(x)}{|x|^2}\,d\mathcal{H}_x^{n-1}\\ \geq \int_{(\partial E)\cap B_1} 2c^2(x)\,|x|^{2\alpha-2}\,d\mathcal{H}_x^{n-1}+\int_{(\partial E)\setminus B_1} 2c^2(x)\,|x|^{2\alpha+2\beta-2}\,d\mathcal{H}_x^{n-1}.\end{aligned}$$

Minimal surfaces

Furthermore, since E is a cone, its curvatures are positively homogeneous of degree -1 and therefore, for all $x \in \mathbb{R}^n \setminus \{0\}$,

$$|c(x)| = \frac{\left|c\left(\frac{x}{|x|}\right)\right|}{|x|} \le \frac{M}{|x|}, \qquad \text{where} \qquad M := \max_{(\partial E)\cap(\partial B_1)} |c|.$$

Minimal surfaces

As a result,

$$\begin{aligned}
&\int_{(\partial E)\cap B_1} 2c^2(x)\,|x|^{2\alpha-2}\,d\mathcal{H}^{n-1}_x + \int_{(\partial E)\setminus B_1} 2c^2(x)\,|x|^{2\alpha+2\beta-2}\,d\mathcal{H}^{n-1}_x \\
\le\;& 2M^2\left[\int_{(\partial E)\cap B_1} |x|^{2\alpha-4}\,d\mathcal{H}^{n-1}_x + \int_{(\partial E)\setminus B_1} |x|^{2\alpha+2\beta-4}\,d\mathcal{H}^{n-1}_x\right] \\
=\;& 2M^2\left[\int_0^1 \rho^{2\alpha-4}\,\mathcal{H}^{n-2}\Big((\partial E)\cap(\partial B_\rho)\Big)\,d\rho + \int_1^{+\infty} \rho^{2\alpha+2\beta-4}\,\mathcal{H}^{n-2}\Big((\partial E)\cap(\partial B_\rho)\Big)\,d\rho\right] \\
=\;& 2M^2\mathcal{H}^{n-2}\Big((\partial E)\cap(\partial B_1)\Big)\left[\int_0^1 \rho^{2\alpha+n-6} + \int_1^{+\infty} \rho^{2\alpha+2\beta+n-6}\,d\rho\right] \\
=\;& 2M^2\mathcal{H}^{n-2}\Big((\partial E)\cap(\partial B_1)\Big)\left[\frac{1}{2\alpha+n-5} - \frac{1}{2\alpha+2\beta+n-5}\right] \\
<\;& +\infty,
\end{aligned}$$

as long as

$$2\alpha+n-5>0 \qquad \text{and} \qquad 2\alpha+2\beta+n-5<0.$$

Minimal surfaces

One can also calculate that

$$|\nabla \varphi_\varepsilon(x)| \leq C \max\{|x|^{\alpha+\beta-1}, |x|^{\alpha-1}\},$$

for some $C > 0$ depending only on α and β.

Minimal surfaces

From these pieces of information, it follows that

$$\begin{aligned}
&\left|\int_{\partial E} c^2(x)\,|\nabla_T\zeta_\varepsilon(x)|^2\,d\mathcal{H}^{n-1}_x-\int_{\partial E} c^2(x)\,|\nabla_T\varphi_\varepsilon(x)|^2\,d\mathcal{H}^{n-1}_x\right|\\
\le\;& C\left(\int_{(\partial E)\cap\left(B_{2\varepsilon}\cup(\mathbb{R}^n\setminus B_{1/\varepsilon})\right)} c^2(x)\,|\nabla\varphi_\varepsilon(x)|^2\,d\mathcal{H}^{n-1}_x+\int_{\partial E} c^2(x)\,|\nabla\tau_\varepsilon(x)|^2\,|\varphi_\varepsilon(x)|^2\,d\mathcal{H}^{n-1}_x\right)\\
\le\;& C\Bigg[\int_{(\partial E)\cap\left(B_{2\varepsilon}\cup(\mathbb{R}^n\setminus B_{1/\varepsilon})\right)} \max\{|x|^{2\alpha+2\beta-4},|x|^{2\alpha-4}\}\,d\mathcal{H}^{n-1}_x\\
&\qquad+\int_{\partial E\cap B_{2\varepsilon}} \varepsilon^{-2}|x|^{2\alpha-2}\,d\mathcal{H}^{n-1}_x+\int_{\partial E\setminus B_{1/\varepsilon}} \varepsilon^{2}|x|^{2\alpha+2\beta-2}\,d\mathcal{H}^{n-1}_x\Bigg]\\
\le\;& C\left(\varepsilon^{2\alpha+n-5}+\varepsilon^{5-2\alpha-2\beta-n}\right),
\end{aligned}$$

which is infinitesimal.

Minimal surfaces

Thus,

$$
\begin{aligned}
&\int_{(\partial E)\cap B_1} 2c^2(x)\,|x|^{2\alpha-2}\,d\mathcal{H}^{n-1}_x+\int_{(\partial E)\setminus B_1} 2c^2(x)\,|x|^{2\alpha+2\beta-2}\,d\mathcal{H}^{n-1}_x\\
\le\;&\liminf_{\varepsilon\searrow 0}\int_{\partial E}\frac{2c^2(x)\,\zeta_\varepsilon^2(x)}{|x|^2}\,d\mathcal{H}^{n-1}_x\\
\le\;&\liminf_{\varepsilon\searrow 0}\int_{\partial E}c^2(x)\,|\nabla_T\zeta_\varepsilon(x)|^2\,d\mathcal{H}^{n-1}_x\\
=\;&\liminf_{\varepsilon\searrow 0}\int_{\partial E}c^2(x)\,|\nabla_T\varphi_\varepsilon(x)|^2\,d\mathcal{H}^{n-1}_x\\
\leqslant\;&\liminf_{\varepsilon\searrow 0}\int_{\partial E}c^2(x)\,|\nabla\varphi_\varepsilon(x)|^2\,d\mathcal{H}^{n-1}_x.
\end{aligned}
$$

Minimal surfaces

And so, by the Dominated Convergence Theorem,

$$\begin{aligned} & \int_{(\partial E)\cap B_1} 2c^2(x)\,|x|^{2\alpha-2}\,d\mathcal{H}^{n-1}_x + \int_{(\partial E)\setminus B_1} 2c^2(x)\,|x|^{2\alpha+2\beta-2}\,d\mathcal{H}^{n-1}_x \\ \le \quad & \int_{\partial E} c^2(x)\,\lim_{\varepsilon\searrow 0}|\nabla\varphi_\varepsilon(x)|^2\,d\mathcal{H}^{n-1}_x \\ = \quad & \int_{(\partial E)\cap B_1} \alpha^2 c^2(x)\,|x|^{2\alpha-2}\,d\mathcal{H}^{n-1}_x + \int_{(\partial E)\setminus B_1} (\alpha+\beta)^2 c^2(x)\,|x|^{2\alpha+2\beta-2}\,d\mathcal{H}^{n-1}_x, \end{aligned}$$

that is

$$\int_{(\partial E)\cap B_1} \kappa_1\,c^2(x)\,|x|^{2\alpha-2}\,d\mathcal{H}^{n-1}_x + \int_{(\partial E)\setminus B_1} \kappa_2\,c^2(x)\,|x|^{2\alpha+2\beta-2}\,d\mathcal{H}^{n-1}_x \le 0$$

where $\kappa_1 := 2-\alpha^2$ and $\kappa_2 := 2-(\alpha+\beta)^2$.

Minimal surfaces

And so, by the Dominated Convergence Theorem,

$$\begin{aligned}
&\int_{(\partial E)\cap B_1} 2c^2(x)\,|x|^{2\alpha-2}\,d\mathcal{H}^{n-1}_x + \int_{(\partial E)\setminus B_1} 2c^2(x)\,|x|^{2\alpha+2\beta-2}\,d\mathcal{H}^{n-1}_x \\
\leq\ &\int_{\partial E} c^2(x)\,\lim_{\varepsilon\searrow 0}|\nabla\varphi_\varepsilon(x)|^2\,d\mathcal{H}^{n-1}_x \\
=\ &\int_{(\partial E)\cap B_1} \alpha^2 c^2(x)\,|x|^{2\alpha-2}\,d\mathcal{H}^{n-1}_x + \int_{(\partial E)\setminus B_1} (\alpha+\beta)^2 c^2(x)\,|x|^{2\alpha+2\beta-2}\,d\mathcal{H}^{n-1}_x,
\end{aligned}$$

that is

$$\int_{(\partial E)\cap B_1} \kappa_1\,c^2(x)\,|x|^{2\alpha-2}\,d\mathcal{H}^{n-1}_x + \int_{(\partial E)\setminus B_1} \kappa_2\,c^2(x)\,|x|^{2\alpha+2\beta-2}\,d\mathcal{H}^{n-1}_x \leq 0$$

where $\kappa_1 := 2-\alpha^2$ and $\kappa_2 := 2-(\alpha+\beta)^2$.

Minimal surfaces

If $3 \leq n \leq 7$ we can choose

$$\alpha := \frac{5-n}{4} + \frac{\sqrt{2}}{2} \qquad \text{and} \qquad \beta := -\sqrt{2},$$

obtain that κ_1 and κ_2 are strictly positive and conclude the proof of version 2!

Minimal surfaces

We also recall that the regularity of minimal surfaces is strictly linked to the so-called Bernstein's problem which asks whether or not a minimal graph in $\mathbb{R}^n$ (i.e., a minimal surface which possesses a global graphical structure of the form $x_n = u(x')$ with $x' \in \mathbb{R}^{n-1}$) is necessarily affine.

The answer to this problem is affirmative in dimension $n \leq 8$ because (by [De Giorgi 1965]) if all minimal cones in $\mathbb{R}^{n-1}$ are halfplanes, then Bernstein's problem has an affirmative answer in $\mathbb{R}^n$.

Minimal surfaces

We also recall that the regularity of minimal surfaces is strictly linked to the so-called Bernstein's problem which asks whether or not a minimal graph in $\mathbb{R}^n$ (i.e., a minimal surface which possesses a global graphical structure of the form $x_n = u(x')$ with $x' \in \mathbb{R}^{n-1}$) is necessarily affine.

The answer to this problem is affirmative in dimension $n \leq 8$ because (by [De Giorgi 1965]) if all minimal cones in $\mathbb{R}^{n-1}$ are halfplanes, then Bernstein's problem has an affirmative answer in $\mathbb{R}^n$.