Coarea formula

Otte Heinävaara

December 5, 2016

1 Coarea formula for linear maps

1.1 Fubini

Coarea formula is a kind of a generalization of Fubini's theorem.

Theorem 1 (Fubini's theorem (in $\mathbb{R}^2$ for Lebesgue measure))**.** *Assume that $f \in L^1(\mathbb{R}^2)$. Then*

$$\int_{\mathbb{R}^2} f(x,y)d(x,y) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x,y)dxdy.$$

This (central) theorem is rather immediate consequence of the following result, a special case of the previous.

Theorem 2 (Fubini's theorem (in $\mathbb{R}^2$ for Lebesgue measure); cheap version)**.** *Assume that $A \in \mathbb{R}^2$ is measureable. Then*

$$m_2(A) = \int_{\mathbb{R}} \int_{R} \chi_A(x,y)dxdy = \int_{\mathbb{R}} m_1(\{x \in \mathbb{R} \mid (x,y) \in A\})dy.$$

The point in both of previous theorems is: if we want to integrate over a region, we can slice the region, integrate over slices, and the them integrate over the results on slices.

1.2 Slicing

We would like to a have generalization for the previous result: what if the slices are not lines parallel to coordinate axis but, say, to the line $x + y = 0$. Now that we are integrating over slices not parallel to coordinate axis, there's a problem. Previously, there was canonical measure on these slices; slices are canonically (under projection) copies of $\mathbb{R}$ itself, so one uses one-dimensional lebesgue measure. One could parametrize the slices, but then again, we have the Hausdorff measure. One should have (again, for measureable set $A \in \mathbb{R}^2$)

$$m_2(A) = \int_{\mathbb{R}} \mathcal{H}^1(\{(x,y) \in A \mid x+y = s\})ds.$$

Note that we integrated here over lines parallel to $x + y = 0$, namely line of the form $x + y = s$ for some s. These line exhaust the plane.

Of course, we could have sliced with lines parallel to some (any) other line, say $3x - 2y = -1$.

Higher dimensional analogue is clear: integrate over hyperplanes parallel to fixed hyperplane, say $x + y + z = 0$ in $\mathbb{R}^3$. The equation reads

$$m_3(A) = \int_{\mathbb{R}} \mathcal{H}^2(\{(x,y,z) \in A \mid x+y+z = s\})ds.$$

But already in $\mathbb{R}^3$ we can do more: we could still slice with lines. Previously, we parametrized the slices with the constant term in the equation of hyperplane but that approach is now awkward. Instead it makes sense to parametrize with points of $\mathbb{R}^2$! Note that even defining lines in $\mathbb{R}^3$ is not as simple as in $\mathbb{R}^2$ but they can be undestood as intersection of two planes. If one wants to slice with lines parallel to the line with parametrization of the form (t,t,t) for $t \in \mathbb{R}$, the line is simply intersection of the planes $x - y = 0$ and $x - z = 0$. This leads to parametrization for all the lines parallel to the previous: they are intersection of planes of the form $x - y = s$ and $x - z = t$ for $s, t \in \mathbb{R}$. Thus, we should have

$$m_3(A) = \int_{\mathbb{R}^2} \mathcal{H}^1(\{(x,y,z) \in A \mid x-y = s, x-z = t\})d(s,t).$$

Similarly for other lines. Also, it doesn't take too much effort to figure out how one could generalize the previous slicing to arbitrary dimension and arbitrary lower dimensional slicings.

1.3 Level sets

Sadly, the previous new formulas aren't quite true: they are only true up to constant. Luckily, it's not too hard to figure out this constant, but we should understand the previous construction bit differently.

In the first case we integrated over lines of the form $x + y = s$. Note that these are just level sets of the map $L : \mathbb{R}^2 \to \mathbb{R}$ given by $L(x,y) = x + y$. The formula might thus more properly (though still off by factor) be written as

$$m_2(A) = \int_{\mathbb{R}} \mathcal{H}^1(A \cap L^{-1}(s))ds.$$

Similarly, for our first $\mathbb{R}^3$ case we should take $L : \mathbb{R}^3 \to \mathbb{R}$ with $L(x, y, z) = x + y + z$, instead. In the final example, we take $L : \mathbb{R}^3 \to \mathbb{R}^2$ with $L(x, y, z) = (x - y, x - z)$ and get

$$m_3(A) = \int_{\mathbb{R}^2} \mathcal{H}^1(A \cap L^{-1}(s, t))d(s, t).$$

More generally, every this kind of linear or affine slicing is defined by a linear map. The codomain of the linear map defines the dimension of the slices: the bigger (i.e. bigger-dimensional) codomain, the smaller (lower dimensional) the slices. The general formula (off by factor) reads

$$m_n(A) = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap L^{-1}(x))dx.$$

Even better, every (non-degenerate, whatever that means) linear mapping defines slicing.

1.4 Coarea formula for linear maps

We still need to determine the constant in the formula: it turns out to be the determinant of the linear map.

Theorem 3 (Coarea formula for linear maps)**.** *For $n \geq m$, measureable $A \in \mathbb{R}^n$ and $L : \mathbb{R}^n \to \mathbb{R}^m$ linear we have*

$$|L| m_n(A) = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap L^{-1}(x))dx,$$

where $|\cdot|$ denotes the (generalized) determinant of L.

Note that we have the determinant on the left-hand side. This is just for convenience. Even if the linear mapping is very degenerate, the right hand side should yield zero, but so does the left-hand side. This choice just means that we don't have to make any special cases.

Also if we just multiply the linear map by constant, c, the left hand side is multiplied c^m but so is the right-hand side, as one easily verifies. Statement is hence plausible. We begin the proof.

Proof. We won't worry about measureability issues; please consult [1] if they really bother you.

We shall invoke the following composition: any linear map $\mathbb{R}^n \to \mathbb{R}^m$ with $n \geq m$ can be decomposed as $S \circ P \circ Q$ where S is symmetric, P is projection and Q is orthogonal. Also $|L| = |S|$ (see [1]).

We will first deal with projections. Their determinant is 1 and the statement reads

$$\begin{aligned} m_n(A) &= \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(L^{-1}(x))dx = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(\{(x', y) \in A \mid x' = x\})dx \\ &= \int_{\mathbb{R}^m} m_{n-m}(\{(x', y) \in A \mid x' = x\})dx. \end{aligned}$$

But this is just Fubini for measure $m_n = m_m \times m_{n-m}$ and χ_A.

In general case we shall invoke the following composition: any linear map $\mathbb{R}^n \to \mathbb{R}^m$ with $n \geq m$ can be decomposed as $S \circ P \circ Q$ where S is symmetric, P is projection and Q is orthogonal. Also $|L| = |S|$ (see [1]).

$$\begin{aligned} |L| m_n(A) &= |L| m_n(Q(A)) \\ &= |L| \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(Q(A) \cap P^{-1}(x))dx \\ &= \int_{\mathbb{R}^m} |S| \mathcal{H}^{n-m}(A \cap Q^{-1} \circ P^{-1}(x))dx \end{aligned}$$

Here we used first the fact that Lebesgue measure is preserved in orthogonal transformation and then the already proven case for projections.

We would still like to somehow transform this to

$$\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap Q^{-1} \circ P^{-1} \circ S^{-1}(x))dx = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap L^{-1}(x))dx.$$

But we have actually managed to drop everything to $\mathbb{R}^m$. Specifically we can use the Area formula! If we apply the Area formula for linear map S and $g = (x \mapsto \mathcal{H}^{n-m}(A \cap Q^{-1} \circ P^{-1} \circ S^{-1}(x)))$ we get

$$\int_{\mathbb{R}^m} |S| \mathcal{H}^{n-m}(A \cap Q^{-1} \circ P^{-1}(x))dx = \int_{\mathbb{R}^m} \sum_{y \in S^{-1}(x)} \mathcal{H}^{n-m}(A \cap Q^{-1} \circ P^{-1}(y))dx.$$

Note that if S is one-to-one, the right-hand-side reduces exactly where we wanted. What if it isn't? Then $|L| = 0$, so also $\dim(L(\mathbb{R}^n)) < m$. It follows that $\mathcal{H}^{n-m}(A \cap L^{-1}(x))$ is 0 for almost every $x \in \mathbb{R}^m$, and the statement is easily seen to hold. $\square$

2 Coarea formula for general Lipschitz maps

In the last section we figured out how to affinely slice integrals. Obvious generalization is to have more general slices. Coming back to $\mathbb{R}^2$ we could slice with concentric circles to get something like

$$m_2(A) = \int_0^\infty \mathcal{H}^1((x, y) \in A \mid x^2 + y^2 = r^2)dr.$$

Again one should think this as follows: consider map $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x, y) = \sqrt{x^2 + y^2}$. We can write the previous as

$$m_2(A) = \int_0^\infty \mathcal{H}^1(A \cap f^{-1}(r))dr.$$

Again, one should worry about constants, and it's not very clear whether the following statement holds. But there's one thing to notice: if f is Lipschitz, it's differentiable almost everywhere, by Rademacher's theorem. Differentiable means almost linear, so f can patched with linear maps. This motivates the guess

$$\int_A Jf(x)dx = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap f^{-1}(x))dx.$$

It turns out that this indeed is the case: this is more general form of Coarea formula. Note that the previous equality is sort of a sum of Coarea formula for linear maps: Jacobian of linear map is just its determinant.

As usual in real analysis, we aim to prove this result by cheating, i.e. cleverly reducing the claim to something much simpler or previously proven, which, in this case, is obviously the linear case.

We will need some lemmas for the proof. First of all, one should naturally check that everything that follows is measureable (at least almost everywhere whenever appropriate), but we will omit this. Main idea is to approximate the Lipschitz function f by linear maps. Of course, this alone is not enough, we want to sort approximate the function with linear maps at countably many sets well enough. That's the main idea; we won't state the lemma's yet, but only after the proof, when we have really figured out what we need.

Theorem 4. *Coarea formula for general Lipschitz maps If $n \geq m$, $A \in \mathbb{R}^n$ is measureable and $f : A \to \mathbb{R}^m$ is Lipschitz, we have*

$$\int_A Jf(x)dm_m(x) = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap f^{-1}(x))dm_m(x).$$

Proof. First, we do some preliminary simplifications. We first reduce the claim to the case where f is everywhere differentiable. By Rademacher's theorem f is already almost everywhere differentiable: if we ignore the set where f is not, the left-hand side doesn't change at all. To be able to say the same thing about the right hand side, we need some estimates for it. In particular, we need the following bound:

$$\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap f^{-1}(x))dm_m(x) \leq C_{n,m} \left(\mathrm{Lip}(f)\right)^n \mathcal{H}^n(A),$$

for some $C_{n,m} < \infty$ depending only on n and m. If we manage to get this, we get rid of the exceptional set of non-differentiability.

Like in the linear case, the interesting case was to consider only non-degenerate maps. That's what we do here too: we assume that $Jf(x) > 0$ everywhere in A. It's not immediately clear why this restriction is justified, but we will deal with the other case later.

As mentioned, the idea is to approximate the Lipschitz map with linear maps: its derivatives. We know that linear maps can be decomposed to symmetric map and a (adjoint) of orthogonal map. Orthogonal map will not have much effect on what's happening, so we try to focus on the symmetric part instead. So question remains: how to fish the symmetric part from the derivative?

We perform the following trick: we extend f (from $\mathbb{R}^n$ to $\mathbb{R}^m$) to map $h : \mathbb{R}^n \to \mathbb{R}^n$, in the following way:

$$\begin{gathered} f = q \circ h, \\ h : \mathbb{R}^n \to \mathbb{R}^n,\ h(x) = (f(x), g(x)) \\ q : \mathbb{R}^n \to \mathbb{R}^m,\ q(x, y) = x. \end{gathered}$$

So pad f with g (to be determined) and only then project. The projection is part is very simple, and if we can make h to have non-zero Jacobian, it will be locally one-to-one, by inverse mapping theorem. The only requirement for g is to satisfy the previous.

Consider the gradient of f.

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

It has m rows and n columns. The condition that the Jacobian of f is non-zero means that the rows are linearly independent. When we extend the f with g we add $n-m$ more rows to get full $n \times n$ matrix and our goal is to get this matrix to be invertible. Actually, that's not very hard. We simply take the components of g to be projections to suitable coordinates: this corresponds to adding a row with single 1 to the matrix. So we are aiming to extend the matrix with rows of the form

$$\begin{bmatrix} 0 & 0 & \cdots 0 & 1 & 0 & \cdots & 0 \end{bmatrix}$$

Note that there are n such rows and they are all linearly independent. It follows that they can't all be in the span of the first m rows (since the span is only m) dimensional, so we may extend matrix with one row and preserve the positivity of the Jacobian. If we keep going like this, we eventually fill the matrix, or equivalently build g with projections.

Note we may only do this construction locally: at some other point same choice rows could lead to singular matrix. That's okay though: there are only finitely many choices, and we may partition the set A according to these choices. To simplify our notation though, let's assume that g is same everywhere.

We are approaching the main argument. We choose sets A_k which (up to set of measure zero) exhaust the set A and in which we can approximate h well with linear mappings. We also want h to be one-to-one in these sets. In set A_k we'd like to squeeze the both sides of the expression close to each other. We allow us a little space, determined by the factor $t > 1$ (same for every k). Then we add the inequalities and let $t \to 1$.

We start manipulating from the right-hand side.

$$\begin{aligned}\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A_k \cap f^{-1}(x))dm_m(x) &= \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A_k \cap h^{-1} \circ q^{-1}(x))dm_m(x) \\ &= \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(h^{-1}\left(h(A_k) \cap q^{-1}(x)\right))dm_m(x).\end{aligned}$$

Here we would like to replace h^{-1} by an the approximating linear mapping: and that's exactly what we do.

$$= \int_{\mathbb{R}^m} \mathcal{H}^{n-m}((h^{-1} \circ S_k) \circ S_k^{-1}\left(h(A_k) \cap q^{-1}(x)\right))dm_m(x).$$

Now the point is that as S_k should approximate h, $h^{-1} \circ S_k$ should behave like identity. We don't need to go that far: only thing we need is that $\mathrm{Lip}(h^{-1} \circ S_k)$ should be not very big, at most t. This also means that $(n-m)$ - dimensional Hausdorff measures of the sets are not increased very much, at most by factor t^{n-m} so we get

$$\begin{aligned}&\leq t^{n-m} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(S_k^{-1}\left(h(A_k) \cap q^{-1}(x)\right))dm_m(x) \\ &= t^{n-m} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(\left(S_k^{-1} \circ h\right)(A_k) \cap (q \circ S_k)^{-1}(x))dm_m(x).\end{aligned}$$

Now we have managed to get rid of the non-linear part of the map f, h!. We are free to use Coarea formula for linear maps and we get

$$= t^{n-m}|q \circ S_k|\mathcal{H}^n(\left(S_k^{-1} \circ h\right)(A_k)).$$

Now we also want $S_k^{-1} \circ h$ to have Lipschitz constant at most t, so

$$= t^{2n-m}|q \circ S_k|\mathcal{H}^n(A_k),$$

and also since S_k should approximate h, $q\circ S_k$ should approximate f, so $|q\circ S_k|\leq t^n Jf(x)$ pointwise. We hence get

$$\leq \quad t^{3n-m}\int_{A_k} Jf(x)dm_m(x).$$

So altogether we have proven that

$$\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A_k\cap f^{-1}(x))dm_m(x)\leq t^{3n-m}\int_{A_k} Jf(x)dm_m(x).$$

In very similar manner we calculate that

$$\begin{aligned}
& \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A_k\cap f^{-1}(x))dm_m(x) \\
= & \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A_k\cap h^{-1}\circ q^{-1}(x))dm_m(x) \\
= & \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(h^{-1}\left(h(A_k)\cap q^{-1}(x)\right))dm_m(x) \\
\geq & \frac{1}{t^{n-m}}\int_{\mathbb{R}^m} \mathcal{H}^{n-m}((S_k^{-1}\circ h)\circ h^{-1}\left(h(A_k)\cap q^{-1}(x)\right))dm_m(x) \\
= & \frac{1}{t^{n-m}}\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(\left(S_k^{-1}\circ h\right)(A_k)\cap (q\circ S_k)^{-1}(x))dm_m(x) \\
= & \frac{1}{t^{n-m}}|q\circ S_k|\mathcal{H}^n(\left(S_k^{-1}\circ h\right)(A_k)) \\
\geq & \frac{1}{t^{2n-m}}|q\circ S_k|\mathcal{H}^n(\left(h^{-1}\circ S_k\right)\circ\left(S_k^{-1}\circ h\right)(A_k)) \\
= & \frac{1}{t^{2n-m}}|q\circ S_k|\mathcal{H}^n(A_k) \\
\geq & \frac{1}{t^{3n-m}}\int_{A_k} Jf(x)dm_m(x),
\end{aligned}$$

as long as we additionally have that $Jf(x)\leq t^n|q\circ S_k|$ in A_k. We hence have chain of inequalities

$$\frac{1}{t^{3n-m}}\int_{A_k} Jf(x)dm_m(x)\leq\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A_k\cap f^{-1}(x))dm_m(x)\leq t^{3n-m}\int_{A_k} Jf(x)dm_m(x).$$

Now sum these inequalities up over all A_k's and let $t\to 1$ and we are done with the first case.

We merely outline the idea for the second case, namely $A\subset\{x\in\mathbb{R}^n\mid Jf(x)=0\}$. More details can be found in [3].

We consider map $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ given by $g(x,z) = f(x) + \varepsilon z$. We can apply Coarea formula for this map and set $A \times B(0,1)$, and note that the left-hand side in the Coarea formula for g is (at most) proportional to ε while right-hand side is essentially the right-hand side for f itself. This squeezes the the right-hand side to 0. $\square$

Lemma 5. *For any measureable $A \subset \mathbb{R}^n$ and Lipschitz mapping $f : A \to \mathbb{R}^m$ we have*

$$\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap f^{-1}(x)) dm_m(x) \leq \frac{\omega_{n-m}\omega_m}{\omega n} Lip(f)^m \mathcal{H}^n(A).$$

Proof (Sketch). For any $j \geq 1$ we may find balls $(B_i^j)_{i=1}^\infty$ such that $\mathrm{diam}(B_i^j) < \frac{1}{j}$ for any $i \geq 1$, $A \subset \bigcup_{i=1}^\infty B_i^j$ and still $\sum_{i=1}^\infty m_n(B_i^j) \leq m_n(A) + \frac{1}{j}$. These covers give rise to covers for the slices. Namely consider functions $g_i^j : \mathbb{R}^m \to \mathbb{R}$ given by

$$g_i^j(y) = \omega_{n-m} \left(\frac{\mathrm{diam}(B_i^j)}{2} \right)^{n-m} \chi_{f(B_i^j)}(y).$$

If one fixes y and j, the balls B_i^j for which $y \in f(B_i^j)$ cover set $A \cap f^{-1}(A)$, and they actually form a $\frac{1}{j}$ cover for that set, so we have by definition

$$\mathcal{H}^{n-m}_{\frac{1}{j}}(A \cap f^{-1}(y)) \leq \sum_{i=1}^n g_i^j(y).$$

Using Fatou's lemma we hence conclude that

$$\begin{aligned}
\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap f^{-1}(x)) dm_m(x) &\leq \int_{\mathbb{R}^m} \lim_{j \to \infty} \mathcal{H}^{n-m}_{\frac{1}{j}}(A \cap f^{-1}(x)) dm_m(x) \\
&\leq \int_{\mathbb{R}^m} \liminf_{j \to \infty} \sum_{i=1}^\infty g_i^j(y) dm_m(x) \\
&\leq \liminf_{j \to \infty} \sum_{i=1}^\infty \int_{\mathbb{R}^m} g_i^j(y) dm_m(x) \\
&\leq \liminf_{j \to \infty} \sum_{i=1}^\infty \omega_{n-m} \left(\frac{\mathrm{diam}(B_i^j)}{2} \right)^{n-m} m_m(f(B_i^j)).
\end{aligned}$$

We then invoke so called isodiametric inequality from which it follows that

$$m_m(f(B_i^j)) \leq \omega_m \left(\frac{\mathrm{diam}(f(B_i^j))}{2} \right)^m,$$

and we get

$$\begin{aligned}
&\leq \liminf_{j\to\infty}\sum_{i=1}^{\infty}\omega_{n-m}\omega_m\left(\frac{\operatorname{diam}(B_i^j)}{2}\right)^{n-m}\left(\frac{\operatorname{diam}(f(B_i^j))}{2}\right)^{m}\\
&\leq \frac{\omega_{n-m}\omega_m}{\omega_n}\operatorname{Lip}(f)^m\liminf_{j\to\infty}\sum_{i=1}^{\infty}\omega_n\left(\frac{\operatorname{diam}(B_i^j)}{2}\right)^{n}\\
&= \frac{\omega_{n-m}\omega_m}{\omega_n}\operatorname{Lip}(f)^m m_n(A),
\end{aligned}$$

our claim. □

Lemma 6. *Let $t > 1$ and $h : A \to \mathbb{R}^n$ Lipschitz such that $Jh(x) > 0$ for any $x \in \mathbb{R}$. Now there exists countable collection of Borel sets, $(D_k)_{k=1}^{\infty}$ symmetric automorphisms $S_k : \mathbb{R}^n \to \mathbb{R}^n$ such that*

- $m_n(A \setminus \cup_{n=1}^{\infty} D_k) = 0$.
- $h\big|_{D_k}$ *is one-to-one for every $k \geq 1$.*
- *For any $k \geq 1$ we have*

$$Lip(S_k^{-1}\circ(h\big|_{D_k})), Lip((h\big|_{D_k})^{-1}\circ S_k) \leq t$$

and

$$t^{-n}|\det(S_k)| \leq Jh\big|_{D_k} \leq t^n|detS_k|$$

Proof (idea). The rough idea is the following: We already had similar lemma for the area formula. There we had essentially same conditions but we had restrictions on the Lipschitz constants of $Lip((h\big|_{D_k})\circ S_k^{-1})$ and $Lip(S_k\circ(h\big|_{D_k})^{-1})$, so h and $S_k's$ had different order. We use first that lemma to simply make our h one-to-one on small sets and then we use the lemma again for the inverse mappings to get our claim. That's pretty much all there is to it. □

References

[1] L. Evans and R. Gariepy, Measure theory and fine properties of functions, CRC Press, 1992

[2] I. Holopainen, Geometric measure theory, 2016.

[3] F. Lin and X. Yang, Geometric Measure Theory: An Introduction, International Press, 2002.