Artif Intell Law (2010) 18:247–283
DOI 10.1007/s10506-010-9097-5

Norm-system revision: theory and application

Audun Stolpe

Published online: 5 October 2010

Abstract This paper generalises classical revision theory of the AGM brand to sets of norms. This is achieved substituting input/output logic for classical logic and tracking the changes. Operations of derogation and amendment—analogues of contraction and revision—are defined and characterised, and the precise relationship between contraction and derogation, on the one hand, and derogation and amendment on the other, is established. It is argued that the notion of derogation, in particular, is a very important analytical tool, and that even core deontic concepts such as that of permission resists a satisfactory analysis without it. By way of illustration the last section of the paper analyses the much debated concept of positive permission, of which there turns out to be more than one kind.

Keywords Derogation · Amendment · Revision theory · Permission

1 Introduction

Formal deontics—the study of norms and normative modalities using mathematical tools—has reasserted itself as a serious academic discipline largely due to its potential for application in computer science. Multi-agent systems, in particular, have been put forth as a principal case, but more generally a deontic level of representation is arguably an essential component of any system that recognises the possibility of failure, non-conformity or deviation from a stipulated optimum (Carmo and Jones 2002). Examples of such systems include on-line marketplaces, disaster response system, virtual supply chains, and e-government services.

There is widespread agreement that a static model of codes of norms runs counter to the dynamic nature of normative systems (see e.g. (Boella et al. 2006, 2009;

A. Stolpe (✉)
Institute of Informatics, University of Oslo, GaustadallTen 23, 0373 Oslo, Norway
e-mail: audus@ifi.uio.no

Carmo and Jones 2002; Governatori and Rotolo 2008; Segerberg 2009)). It is an essential premise at least for *open* multi-agent system that the behaviour and the interaction of agents in the system cannot be accurately predicted. The system may therefore need to adapt to changing circumstances, new norms may need to be created and old ones removed. The system may also need to utilise different subsets of norms under different circumstances, as for instance when an agent violates a primary obligation, thereby bringing a reparational norm into play, or it may need to choose which one of two conflicting norms that should prevail for a given case in which there is a conflict of norms (see e.g. (Hansen 2004)).

Indeed, it can be argued that several core deontic concepts are themselves inherently dynamic, and that an account of a code's modes of transformation is one of the tools needed to characterise those concepts. Consider for instance permission. It is commonly agreed that permitted actions fall under one of two broad kinds; those that are *negatively* permitted and those that are *positively* permitted or permitted in decree. Negative permission is simple, and can be regarded as the dual of obligation just as possibility is the dual of necessity in alethic modal logic. Positive permission, on the other hand, is more elusive. A permissive legal norm—regulating, say, the collection and dissemination of personal information—is often formulated along the following lines: "Personal information may only be processed by the consent of the registered person, or if processing is statutorily warranted, or if such processing is required in order to honour an agreement with the registered person (…)". The word 'only' in the opening sentence indicates that processing of personal information is by default prohibited. The text then goes on to list a set of explicitly recognised cases that are excepted from the ban and which therefore constitute permitted actions. Permission is implemented, one might say, by restricting the set of norms that are deemed applicable under specified circumstances: There is a general norm that is deemed to be valid or applicable *by default* and always, save for the cases that are subsumed by the explicitly stated permissions. When one of the latter cases is satisfied, the set of applicable norms contracts to exclude the general prohibition, leaving a smaller code in force. This, I shall argue, is at least a fruitful way to look at it.

I shall refer to the above mentioned contraction operation as *derogation*. In law the term 'derogation' denotes the partial repeal or abrogation of a law by a later act that limits its scope or impairs its utility and force. For example, statutes in derogation of common law are those statutes which effect a change in the principles of precedent developed in earlier case law. Abstracting away the temporal relationship between the derogans and the derogandum as well as the sources of law (the issuing authorities), derogating from a norm n means limiting the scope of n by registering *exceptions* to it. It is this more general notion that is the principal target of analysis in the present paper. I take it to be sufficiently close to the legal sense of 'derogation' to warrant a reuse of the term. The reader should bear in mind, though, that the legal concept is really a special case of the notion analysed here.

Derogation, so understood, will be analysed in terms of *contractions* on a set of norms, which is a generalisation of the corresponding notion in the so-called AGM-framework of theory revision (Alchourron and Makinson 1982; Alchourron et al. 1985). Since its inception in the 1980s, revision theory has grown into a wellstudied research programme that branches off into numerous other disciplines, such as epistemology, artificial intelligence and multi-agent systems. Two of the main strengths of the idiom is its generality and its wide applicability (Williams and Rott 2001). Change, considered *in abstracto*, is reduced to consistency-preserving addition and subtraction of items to and from a set of formulae modulo some notion of implicature. This 'beating heart' perspective on change is simple and compelling, and has turned out to be a robust and sustainable mathematical idiom.

The reader may be more used to thinking of exceptions in terms of non-monotonic inference rather than in terms of contraction. From a philosophical point of view it is of little consequence which idiom we choose, however, for it is well known that theory revision and non-monotonic inference are two sides of the same coin: Every operator of (AGM) contraction determines a revision operator via the Levi identity, and every revision operator may be considered a non-monotonic consequence operator when projected onto its right argument (the two-place operation of revising K with a, written $K{*}a$, becomes the unary operation of *revision-with-a modulo K*, written $*_K$ (a)). Conversely, any non-monotonic consequence operation can be seen as a revision function, which in turn determines a contraction function via the Harper identity (see. Makinson 2003, chapter 6; Makinson and Gärdenfors 1991). Therefore, in the principal case of revision where K is contracted so as to exclude $\neg a$ prior to the addition of a, one may if one likes equally well speak of a as an *exception* to the default assumption $\neg a$ implicit in K.

Now, in order to move towards a general theory of norm-system dynamics, similar in general principle to the AGM paradigm, what we need first of all is a well-defined notion of a *code* or *system* of norms, construed in analogy to the notion of the Tarski-closure of a set of sentences. A closure operator that fits the bill, or so I shall argue, is supplied by input/output logic as set out in a series of papers by Makinson and van der Torre (2000, 2001, 2003a, b). Indeed AGM revision and input/output logic turn out to combine beautifully to give us exactly what we need in order to move towards a general theory of norm system dynamics. I shall attempt to illustrate the utility of the resulting theory—for purposes of conceptual analysis—towards the end of the paper, by analysing the concept of positive as outlined above.

The layout of this paper is as follows: Sect. 2 gives an informal introduction to the input/output idiom, with a special emphasis on its relation to the philosophy of norms. Sect. 3 records some logical properties of the theory of simple-minded output, so-called, that will be used as lemmas in subsequent sections. It focuses on three notions in particular; equivalence, complementation and maximal consistent subsets—all modulo input/output implicature. Section 4 gives a complete AGM-style characterisation of the derogation operation, whilst Sect. 4.1 provides a mapping into classical AGM contraction that can be used to migrate results from classical contraction theory over into the theory of derogation and vice versa. Norm-system *revision*, or *amendment* as I shall call it, is defined and characterised in Sect. 5. The paper ends with a case study in Sect. 6 of the concept of positive permission, designed to illustrate the utility of the derogation operator for purposes of conceptual analysis.

The reader should note that there is some literature on the topic of norm-system dynamics already. Independently of the present work Boella et al. (2009) proposed that a theory of norm-system dynamics be built precisely on AGM revision and input/output logic. Boella et al. fix attention on a set of axioms (or 'postulates') that are considered plausible candidates for an operator of norm-system contraction. However, it is a weakness of their account that no construction is offered against which the adequacy of these postulates can be measured. That is, Boella et al. do not map their postulates to a semantics (broadly understood) that gives the meaning of the postulates in terms of operations on the code of norms. As a consequence, certain subtleties of input/output logic escape notice in the formulation of the postulates. I shall comment on some of these points along the way. Other recent sources that discuss norm-system dynamics include Governatori et al. (2005, 2007, 2008). Here the focus is on the notions of legal abrogation and annulment, and the formalism is a temporal and modal extension of defeasible logic. It is thus different from the present paper with respect to the underlying logic. Moreover the central concern seems to be to capture a notion of retroactive legal change, rather than, as in the present paper, to develop an abstract and general model of norm-system evolution.

2 Input/output logic as a logic of norms

A few notational preliminaries first: We use lower case letters a, b, c … to range over formulae of classical propositional logic, denoted L. The distinguished letters t and f will stand for arbitrary tautologies and contradictions respectively. Sets of formulae are denoted by upper case letters from A to D. Upper case letters from F to I denote sets of norms, that is, binary relations over subsets of L. When $G \subseteq L \times L$ we denote its pre-image under L as G^1 and its image under L as G^2. Image-formation will be denoted by ordinary parentheses, for instance $(G \cup H)(a)$ denotes the image of the relation $G \cup H$ under a. Classical consequence is written with a turnstile $\vdash$ when considered as a relation over $2^L \times L$, and as Cn when viewed as an operation on 2^L onto itself. To make the notation less verbose, we follow the convention of writing $A \cup a$ instead of $A \cup \{a\}$, and similarly for norms.

Just as the theoretical paradigm of a *theory* is a logically closed set of sentences (i.e. a set of sentences closed under entailment), the theoretical paradigm of a normative system may be taken to be a set of *prima facie* mandatory norms that contains all norms it entails. In input/output logic a (*prima facie*) norm is simply a pair (a, b) correlating an *applicability condition*, *trigger* or *input* a with a *duty*, *optimality condition* or *output* b—these will sometimes be denoted neutrally as the antecedent and consequent of a norm respectively. The correlation between the antecedent and consequent is *logically* arbitrary in the sense that a pair is not a formula, so there is nothing to the norm (a, b) over and above the fact that some authority requires that b be done given a. One could see this as an expression of a kind of anti-naturalism, or conventionalism, w.r.t. to norms. The validity of a norm (a, b) need not have any ontological or epistemological status beyond that of being decreed to hold. A *code* of norms in input/output logic is simply a set G of such pairs, from which it follows that the explicitly declared requirements, in any situation a (or alternatively, on any *input* a) according to G, can be obtained by taking the image of G under a. The fundamental notion of normative implicature in turn allows implicit norms to be derived from the explicit ones—i.e. from the ones contained in G—e. g. by recognising that which implies (logically) the trigger of a norm as itself a trigger of a norm, and that which follows (logically) from an explicitly declared requirement as itself mandated by a norm. To be more precise, the simplest model of a normative system is an operation *out* of type $2^{L \times L} \times L \mapsto 2^L$ defined as follows:

Definition 1 $out(G, a) = Cn(G(Cn(a)))$.

I shall allow myself to represent this *out*-operator sometimes by projecting it onto its left argument, writing $out(G)$. These two modes of expression will be assumed equivalent by putting $(a, b) \in out(G)$ iff $b \in out(G, a)$.

Input/output logic, of which definition 1 gives the simplest example, addresses a fundamental problem in the logic of norms independently observed by Dubislav (1937), and Jørgensen (1937): On the one hand, truth-functional operators like ‘and’, ‘or’ and ‘not’ are routinely applied to items construed as norms, forming complex norms out of primitive ones. On the other hand, norms are not in general *true or false*, i.e. they do not in general have truth-values, so it is not clear what could be meant by such compounding (Makinson 1999). To be sure, norms are clearly content-full in *some* sense and perhaps *some* norms have truth values, for instance moral norms, as the moral realist claims. He may well be correct that it simply is wrong to take pleasure in another’s pain, to taunt and threaten the vulnerable, to prosecute those known to be innocent, to torture babies for fun, and to sell another’s secrets solely for personal gain (Shafer-Landau 2003, p. 248). Perhaps one need not infer such principles from other beliefs in order to be justified in holding them true, perhaps we should just accept them as self-evident and ask for no further argument. After all, there are other kinds of statements we would be willing to accept without argument, for instance that an object cannot be both red and white all over, or, that if you are suffering then you know that you are suffering. In such cases, our belief is direct and immediate, and if asked for reasons we tend to reply ‘don’t you just see it?’, without being able to give any deeper or more fundamental reason than that. Perhaps specifically moral norms are true in this sense. If that is the case, then a valid moral argument is one that preserves truth, just as logically valid arguments do. Ethical controversy, then, is primarily about truth, and the authority of moral norms is proportional to the legitimacy of their truth-claim [ibid.,30].

Yet, it should be obvious that this property, if it exists, does not generalise. Consider for instance the regulation from Norwegian law that requires all shops to be closed on Sundays except those that deal mainly in groceries and have a total sales area of less then a hundred square meters. It would certainly require an ingenious argument to convince people that this norm is true *as such*, that is, true as a *stance-independent* feature of reality. It is much more natural to view it as a *logically* arbitrary stipulation laid down by the Norwegian authorities for political reasons.

Of course one could say that it is true *of* Norwegian law that the regulation is part of *it*. This route has often been chosen by norm-theorists of a formal bent, who continue to rely on truth to supply a notion of inference. The idea in general outline, familiar I am sure to many, is that the Jörgensen/Dubislav problem may be solved by making a distinction between two uses of norm-sentences; norm sentences can be used for *prescriptive* purposes to influence and direct the behaviour of agents, or it can be used descriptively to state that something is obligatory, permitted or prohibited *according to a given system of norms*. Consider the following example from (Hilpinen 2001):

$$\text{Motor vehicles ought to use the right-hand side of the road.} \quad (1)$$

Such a statement can be used or uttered *performatively* to direct the behaviour of the citizens of a given country, say Norway, but it can also be used to state something *about* Norwegian law, which is *true* of *it*, but false of other legal systems such as e.g. that of the UK. It is true *of* Norwegian law, but false with respect to British law, that such a norm belongs to it. I take it that the distinction between descriptive and prescriptive uses of a norm needs no lengthy elaboration. For the sake of precision, we may put the point in terms of assent. If I assent, sincerely, to a sentence attributing a norm to a code or normative system, then I form a corresponding belief about that system. If, on the other hand, I assent to a norm used prescriptively—say one that requires me to drive on the right hand side of the road—then I cannot be assenting sincerely if I do not form an *intention* to act accordingly.

In the philosophical literature descriptive norm sentences are usually called normative statements or norm-propositions. I prefer 'norm-attributions' myself; what such statements do is essentially to assert the *membership* of norms in the sets that constitute the system. Now, norm-attributions unlike norms simpliciter *are* true or false of course. The logical relations between them can therefore be understood in the usual way in terms of truth-preserving inferences. This has been perceived by many to put us in a position to read off systematic relations between the norms *themselves* from the logical relations holding between the norm-attributions. Such a view has been championed by e.g. Alchourrón and Bulygin (1981), von Wright (1998) and by Hans Kelsen in his classic *Pure Theory of Law*:

> Since legal norms, being prescriptions … can neither be true nor false, the question arises: How can logical principles, especially the Principle of the Exclusion of Contradictories and the Rules of Inference be applied to the relation between legal norms, if according to traditional views these principles are applicable only to assertions that can be true or false. The answer is: Logical principles are applicable, indirectly, to legal norms to the extent that they are applicable to the rules of law which describe the legal norms and which can be true or false … one norm may be deduced from another if the rules of law that describe them can be deduced from a logical syllogism ((Kelsen 1967), p. 74).

What Kelsen calls 'rules of law' are just norm-attributions, and Kelsen takes them to provide an indirect route to the logic of norms proper. This claim faces one basic difficulty from which there is, in my opinion, no recovery; it is either trivial or viciously circular. This becomes clear as soon as one asks oneself 'what determines the membership of the norms in the system?'. Of course, the answer to that depends on how we construe the concept of a system. We have essentially only two choices: We may say that *any* set of norms G is a normative system. In that case the only norm-attributions we can form are purely boolean combinations of simple membership assertions 'it is true of Norwegian law that wearing a mask in public is in general prohibited *and* it is true that driving on the right is mandatory' and so forth. Alas, if this is all we can do, then there simply is no logic of norms or normative reasoning beyond boolean logic. The other option is to say that the system is the *closure* $C(G)$ of G under some suitable operation C. Depending on C this could perhaps give us more interesting and non-trivial norm-attributions. Clearly, though, we would have to know the behaviour of C *before* we can know which norm-attributions are true of $C(G)$. Thus C, whatever it is, must be construed as *conceptually prior* to the norm-attributions. Stated differently, the truth of norm-attributions can only be determined if we already know the logic that governs the norms, not vice versa. Norm-attributionists have the tail wag the dog, therefore. True, any given norm n of a normative system s can be mapped to a corresponding statement ' 'n' is a valid norm of s', but we come to know that that is true by deriving n from s.

Judging by the forerunner (Makinson 1999) as well as the original input/output papers themselves (Makinson and van der Torre 2000, 2001, 2003), the Jørgensen/Dubislav-problem was the principal motivation for the introduction of the input/output idiom. Input/output logic is conceived as a logic of norms from its very inception:

> Input/output logic takes its origin in the study of conditional norms. These may express obligations under some legal, moral or practical code, goals, contingency plans, advice, and so on. They may be expressed in imperative form, in such-and-such a situation, do so-and-so, or in indicative form, in terms like In such-and-such a situation, so-and-so should be the case, or … should be brought about, or … should be worked towards, or … should be followed—these locutions corresponding roughly to the kinds of norm mentioned (Makinson and van der Torre 2003, p. 1).

In other words, input/output attempts to tackle the Jørgensen/Dubislav-problem head-on by offering an alternative to the proposition as the paradigm of a norm. The proposition is replaced with an abstract conception of a norm whose principal characteristic is that a conditional norm is no longer construed as a *conditional* in the logical sense, and is not assumed to have a truth-value. More specifically, a norm is taken to be a logically arbitrary stipulation and the role of logic is seen as a modest one: To preprocess input before it goes in to the system and to unpack the output on the other side (Makinson and van der Torre 2003).

The *out* operator from definition 1 admits a syntactical representation in terms of a system of inference rules:

Theorem 1 *For any G, $out(G)$ is precisely the set of norms that are derivable in the system consisting of axioms (t, t) and (a, b) for all $(a, b) \in G$, together with the inference rules (this is [Observation 1,38]):*

$$SI\ \frac{(c,b)}{(a,b)}\ \ if\ a \vdash c \qquad AND\ \frac{(a,b),(a,c)}{(a,b \wedge c)} \qquad WO\ \frac{(a,b)}{(a,c)}\ \ if\ b \vdash c$$

Here as in (Makinson and van der Torre 2003) a *derivation* is understood in terms of rules as follows: A *rule* r of arity $n \geq 0$ is an $n + 1$-ary relation over the set $L \times L$ of pairs of formulae in the language L. For any element $((a_1, b_1), \ldots, (a_n, b_n), (a_{n+1}, b_{n+1})) \in r$ we call $(a_1, b_1), \ldots, (a_n, b_n)$ the premises of the rule and (a_{n+1}, b_{n+1}) its conclusion. A derivation of a pair (a, b) from a set G of pairs of formulae, given a set R of rules, is understood to be a tree with (a, b) at the root, each non-leaf node related to its immediate parents by the inverse of a rule in R, and each leaf node either the conclusion of a zero-premise rule in R, or an element of G, or of the form (t, t). If we denote the set of elements that are derivable from G in the system above as $deriv(G)$ then theorem 1 states that $out(G) = deriv(G)$. Note that these operators are closure operators, that is, they satisfy monotony, inclusion and idempotence.

The reader should note that the *out*-operator considered here corresponds to the one Makinson and van der Torre call simple-minded output (2000)—it is the least comitting of the input/output operators. Although quite weak, it has the great virtue that it can be viewed as a most natural and immediate generalisation of classical logic. More precisely, classical logic is the special case where the set of norms G is the diagonal relation over L:

Theorem 2 *Let G be the diagonal relation over L. Then* $out(G, a) = Cn(a)$.

Proof We prove only the left-to-right-direction, the converse is trivial: Suppose $b \in out(G, a)$. By compactness for classical consequence there is a set $(a_1, b_1), \ldots, (a_n, b_n) \in G$ such that $a \vdash a_j$ for $1 \leq j \leq n$ and $\bigwedge_{i=1}^{n} b_i \vdash b$. Since G is the diagonal over L we have that $a_j = b_j$ for $1 \leq j \leq n$, whence $\bigwedge_{i=1}^{n} a_i \vdash b$. Thus, $a \vdash \bigwedge_{i=1}^{n} a_i \vdash b$ so $b \in Cn(a)$ as desired. □

Thus, simple-minded output makes a small step up the ladder of abstraction, substituting ordered pairs for conditional statements. This has two effects: First, outputs cannot be recycled as inputs, meaning that norms cannot be chained together. Chaining of conditional obligation is a contentious issue (see e.g. Hansen 2004; Stolpe 2008b; von Wright 1998; Vranas 2008) and it is by no means obvious that normative implicature is transitive. There are ways to add transitivity on top of simple-minded output (the reader is referred to the original input/output papers), but it should be considered a feature, not a bug, that it is not integral to the idiom. Secondly, conditional norms are not in the general case reflexive. This is very important, and captures an essential feature of normative reasoning: The whole point of having a logic of norms is to represent and reason about the potential *discrepancy* between the actual and the ideal. If conditional norms are taken to be reflexive, the actual is always ideal so the distinction collapses. Non-reflexivity can be seen as a characteristic feature of all forms of *conventional generation*. Whatever holds according to a norm holds as a matter of conventionally accepted implicature. Indeed there may be warrant for speaking about *non-reflexive logics* as on a par with *non-monotonic logics*. The latter denotes a broad range of inference patterns that in some way or other involve assumptions of falsity, which in turn make arguments sensitive to the absence of information. The former may in a similar spirit be taken to stand for a family of inference patterns that share the characteristic that they are mediated by convention. Examples include permissive norms, mandatory norms, norms of etiquette and constitutive norms (which should also not be construed as reflexive). Simple-minded output has the virtue that it isolates this feature in its purest form, and makes it easy to track the things that change. Of course, the operative assumption is that we will learn something useful that we can apply to more sophisticated versions later.

Having said that, the reader may already find reason to pause at the principles that simple-minded output *does* satisfy, in particular the principle of input strengthening (*SI*). Uncontroversially, normative reasoning is not monotonic, so input strengthening seems counter-intuitive. The thing to keep in mind though is that simple-minded output constitutes a model of *prima facie* norms only. The whole point of the present paper is to investigate ways of overriding or suspending such norms. Input strengthening is *prima facie* valid. In the absence of evidence to the contrary an obligation that applies to one context also applies to stronger contexts, but if contrary information can be brought to bear then that inference is blocked. Thus input strengthening is not simply discarded. When it fails it fails instructively with reference to a particular context. This is as it should be.

3 Properties of the theory of simple-minded output

3.1 Complementation

The conditions under which a normative system ought to be counted inconsistent is a topic of much debate in the literature, and no clear consensus seems to have emerged. Yet, this is an issue pertinent to the matter at hand, since revision—the topic of Sect. 5—is usually defined in terms of it. I shall later follow suit and define the revision of a set of norms G with a norm (a, b) as the removal of *incompatible* material from the consequences of G prior to the addition of (a, b). It is therefore necessary to take a stand on what ‘contradiction’ should be taken to mean when applied to norms, and whether ‘incompatible material’ should be understood in this sense.

In von Wright (1963), Kelsen (1967), takes the coexistence of conflicting norms from the same source to provide an adequate criterion for the contradictoriness of a code of norms: The giving of two conflicting norms is the expression of an irrational will; it is a performative self-contradiction and as such a pure fact that fails to create a norm (Hansen 2004, p. 486). Kelsen says: “To say that *a* ought to be and at the same time ought not to be is just as meaningless as to say that *a* is and at the same time that it is not. A conflict of norms is just as meaningless as a logical contradiction” (Kelsen 1967, p. 205–206). In Alchourrón and Bulygin (1971, p. 62) express the same view: “Generally speaking, a normative system α is *inconsistent* in a case C_i (…) if α correlates C_i with two or more solutions in such a way that the conjunction of these solutions is a deontic contradiction". Translated into input/output idiom this criterion deems a code G inconsistent if $out(G)$ contains (a, f) for some a.

It is remarkable that all the above mentioned writers later reject this view. Kelsen (1991) ultimately comes to view logic as simply inapplicable to norms, and therefore to law. C.E. Alchourrón and E. Bulygin argue that a system of norms that it is impossible to obey might be unreasonable and its norm-giver blameworthy, but its existence does not constitute a logical contradiction [ibid.,20]. Similarly, in (1999) von Wright concedes that existing normative systems may or may not be contradiction-free, and reformulates deontic principles as meta-norms for consistent norm-giving [ibid.,20].

All continue, to regard the coexistence of conflicting norms as a flaw, though. Yet, if we regard a normative system as a set of norms that is closed under some notion of entailment, then there are good reasons, why a system that contains (a, f) need not be deemed neither inconsistent nor incoherent. It is common for a code of norms—probably unavoidable if the code is of any complexity—to regulate two distinct states of affairs separately, which may nevertheless happen to occur simultaneously. Whether or not two events can occur at the same time, is usually a matter of contingent fact, and not a possibility that can be ruled out by stipulation. Say you have volunteered for the local fire brigade and also for the rescue service in the local Red Cross unit. In case of an avalanche you are obliged to assist the Red Cross on site, and similarly, in case of a fire you are obliged to assist the fire brigade at *that* site. Since it is clearly possible that a fire breaks out at the same time as an avalanche occurs, and since it is possible for these events to occur at locations far removed from one another, then in a not too far-fetched scenario you may happen to have taken upon you two incompatible commitments. It is nevertheless clearly not absurd or practically irrational, in any sense we can make of it, for you to volunteer both for the Red Cross and for the fire brigade. That an agent cannot, under unhappy circumstances, live up to all his commitments, is something a code should give us the resources to *anticipate*, it is not something the code should seek to prevent. Indeed, one of the things we should expect a formal model of normative systems to do is precisely to alert us to such conflicts, and the circumstances under which they arise. *Derived* norms (a, f), at least, ought to be read this way. According to this interpretation the derivability of (a, f) is a feature, not a bug, and the norm (a, f) can be regarded as intuitively equivalent with the 'contrapositive' $(t, \neg a)$—the general prohibition against doing a.[1]

In order to talk about these issues with a reasonable level of precision, we may call the norm $(a, \neg b)$ the *local negation* of (a, b), and the norm $(\neg a, \neg b)$ the *global negation* of (a, b). From a lattice-theoretic point of view it is the global negation that is the complement of the norm. Consider the operations $\veebar$ and $\barwedge$ defined as follows:

[1] The inference from (a, f) to $(t, \neg a)$ is not, however, licensed by any of the input/output systems currently on offer.

$$(a,b)\underline{\vee}(c,d) := (a \wedge c, b \vee d)$$

$$(a,b)\overline{\wedge}(c,d) := (a \vee c, b \wedge d)$$

Since the operations $\overline{\wedge}$ and $\underline{\vee}$ are defined by way of boolean conjunction and disjunction the lattice equalities linking conjunction and disjunction are reflected by $\underline{\vee}$ and $\overline{\wedge}$. Take the principle of absorption $a \vee (b \wedge a) = a$ as an example:

$$\begin{aligned}(a,b)\underline{\vee}((a,b)\overline{\wedge}(c,d)) &= (a,b)\underline{\vee}(a \vee c, b \wedge d)\\ &= (a \wedge (a \vee c), b \vee (b \wedge d))\\ &= (a,b)\end{aligned}$$

All the other lattice equalities can be reproduced in a similar manner, and thus $\langle L^2, \underline{\vee}, \overline{\wedge}\rangle$ is a distributive lattice. Now, for any $(a, b) \in L^2$ we have:

$$(f,t)\underline{\vee}(a,b) = (a \wedge f, b \vee t) = (f,t)$$

and

$$(t,f)\overline{\wedge}(a,b) = (t \vee a, f \wedge b) = (t,f)$$

Thus (f, t) is the top and (t, f) the bottom element of the lattice (or, in the infinite case, of its congruence). Now, since

$$(a,b)\overline{\wedge}(a,b)' = (a \vee \neg a, b \wedge \neg b) = (t,f) = \bot$$

it follows that $(\neg a, \neg b)$ is the lattice-theoretic complement of (a, b). This complement is unique, since $\langle L^2, \underline{\vee}, \overline{\wedge}\rangle$ is a distributive lattice (Davey and Priestley 2002, chapter 4.13).

This strongly suggests treating G as inconsistent iff $(t,f) \notin out(G)$. We can then derive the welcome property that an inconsistent set of norms coincides with the total relation L^2 over L:[2]

Theorem 3 $(t, f) \in out(G)$ iff $out(G) = L^2$.

Hence, just as a set of sentences is consistent iff it does not contain the entire language, a set of norms, according to the definition above, is consistent iff it does not contain the total relation.

This criterion—the inclusion of (t, f) in the system—is similar to one proposed by Vranas in (2008) according to which a set of norms is inconsistent iff it contains an *omniviolable* norm, that is, if it contains a necessarily violated norm. Conditional norms with a consistent condition of application are not omniviolable in this sense, since the applicability condition may not be satisfied. If it is not, then the obligation is neither violated nor fulfilled—it simply doesn't apply. It seems more reasonable to say, as Vranas does, that the obligation is *avoided*. Stated differently, even if $out(G)$ contains (a, f), a may not be true whence f may not be derivable from $out(G)$ *under the prevailing circumstances*. Hence, the system may continue to give consistent directions for other cases than a, so it is not *on the whole* inconsistent. It is just a system that flags a as a source of problems—a state of affairs to be avoided.

[2] Proofs of the lemmas and theorems in this subsection—they are all very easy—can be found in Stolpe (2008a).

Truly inconsistent is the system that yields *f always*, for *all* states of affairs. The only norm that does that, the only *omniviolable* norm, is (t, f).

This should make it fairly obvious that overall inconsistency is not the right notion to focus on for purposes of norm-system revision, because the aim of revising G with (a, b) is to make b consistently applicable *to the context described by a*. A removal of (t, f) from G does not suffice for this purpose, since it is entirely possible to have $(t,f) \notin out(G)$ whilst $(a, f) \in out(G)$ for some a. Removing (f, f), on the other hand, would restore consistency within a, but at the cost of purging inconsistencies from *all* contexts. But it has just been argued that one essential service that a model of a normative systems should provide, is the ability to diagnose potential sources of problems by predicting e.g. that (a, f). What we want, therefore, is just local consistency—we want to purge the context a of conflicting elements, but we do not want the process to spill over into unrelated contexts.

As, the next lemma shows, removing locally inconsistent elements will suffice for overall *local* consistency:

Theorem 4 *For any* $a \in L$, $\langle a \times L, \bar{\wedge}, \veebar \rangle$ *is a sub-algebra of* $\langle L^2, \bar{\wedge}, \veebar \rangle$, *and local negation isthe unique complement.*

Thus the local negation of (a, b) is the locally unique complement in the sub-lattice of L^2 determined by a, whence revision with local negation always restores (exempt a few limiting cases) consistency to the context expressed by the antecedent of the norm that is scheduled for addition.

Note that local negation requires attention to limiting cases, for as the next example shows it does not preserve equivalence modulo *out*:

Example 1 Consider the two norms (a, t) and (b, t) and assume that a and b are elementary letters. Both norms have tautologous consequents, and are therefore equivalent modulo *out*. Their local negations are (a, f) and (b, f). Since a and b are logically independent it follows that $(a,f) \notin out((b,f))$, so (a, f) and (b, f) are not equivalent.

As we shall see in Sect. 5, the non-preservation of equivalence under local negation affects the revision operator which will be defined via the Levi identity as the removal of the local negation of the element to be added. Special provisos must therefore sometimes be introduced to hedge against limiting cases. These provisos take the form of appeals to one of the following properties that give sufficient conditions under which preservation of equivalence *does* hold:

Lemma 1 *If* $out((a, b)) = out((c, d))$ *and* $\nvdash b$ *then* $b \equiv d$ and $a \equiv c$.

Proof Suppose $out((a, b)) = out((c, d))$. Then $(a, b) \in out((c, d))$. Since $\nvdash b$, (a, b) must be derived from (c, d) using *SI*, *WO* or both. It follows that $a \vdash c$ and $d \vdash b$. Conversely, we have $(c, d) \in out((a, b))$. Moreover $\nvdash d$, since $d \vdash b$ and $\nvdash b$. Hence (c, d) must be derived from (a, b), whence $c \vdash a$ and $b \vdash d$. □

Lemma 2 *If* $out((a, b)) = out((c, d))$ *and* $a \equiv c$ *then* $b \equiv d$.

Proof Suppose that b is not equivalent to d and that $a \equiv c$. Then we have $out((c,d),a) = Cn(d) \neq Cn(b) = out((a,b),a)$, so $out((c, d)) \neq out((a, b))$. □

3.2 Deduction properties

Since norm-system revision, or amendment as it shall be called, will be analysed in terms of consistency preserving *additions*—this is Sect. 5—it will be expedient to establish a few results that has to do with reasoning about a code in the presence of additional rules. In classical logic reasoning in the presence of extra *assumptions* is summed up by the *deduction theorem* which says that a proposition q can be derived from K in the presence of an extra assumption p iff the proposition $p \rightarrow q$ can be derived from K directly. A similar property holds for norms, and it will be utilised extensively in subsequent sections (proofs can be found in Stolpe (2008a, chap. 4). We have

Lemma 3 (Easy half of deduction) *If* $(a \wedge c, b \rightarrow d) \in out(G)$ *and* $c \vdash a$ *then* $(c, d) \in out(G \cup (a, b))$.

and the converse

Lemma 4 (Hard half of deduction) *If* $(c, d) \in out(G \cup (a, b))$ *then* $(a \wedge c, b \rightarrow d) \in out(G)$.

In the limiting case where $(c, d) \in G$ already (a, b) need not be required for its derivation. Hence, there need not be any logical relationship between a and c. However, in all other cases—that is in all cases where (a, b) is a leaf in the derivation of (c, d) from G—we have that $c \vdash a$, and may thus conclude that $(c, b \rightarrow d) \in out(G)$. To see this, note first that the antecedent of the root of a derivation will be logically stronger than the body of any other rule closer to the leaves:

Lemma 5 *For each node* n: (a, b) *in a derivation, write* L_n *for the set of leaves in the subtree determined by* n. *Then, if* $(c, d) \in L_n$ *then* $a \vdash c$.

Proof By induction on n and the last rule in a derivation. Let n: (a, b) be any node in the tree and suppose $(c, d) \in L_n$. For the base case we have: If (a, b) is a leaf of the tree, then $(a, b) = (c, d)$ so $c \vdash a$ as desired. For *SI*, suppose (a, b) is derived from p: (g, h) by *SI*, then $a \vdash g$. By the induction hypothesis the property holds for p, and thus $g \vdash c$. Hence $a \vdash c$ as desired. For *WO*, suppose (a, b) is derived from p: (g, h) by *WO*. Then $a \equiv g$. By the induction hypothesis we have $g \vdash c$, so $a \vdash c$ as desired. For *AND*, suppose (a, b) is derived from p: (g, h) and q: (g', h') by *AND*. Then $g \equiv g' \equiv a$. By the induction hypothesis we have $g \equiv g' \vdash c$ so $a \vdash c$ as desired. □

The following input-entailment property falls off immediately:

Lemma 6 *If* $(c, d) \in out(F \cup (a, b)) \setminus out(F)$ *then* $c \vdash a$.

Proof Since $(c, d) \in out(F \cup (a, b)) = deriv(F \cup (a, b))$ but not in $out(F) = deriv(F)$. It follows that (a, b) is a leaf in some derivation of (c, d) from $F \cup (a, b)$. Hence $c \vdash a$ by lemma 5. □

I shall have occasion to appeal to this later.

3.3 Remainders

In classical revision-theory, a contraction on a set is carried out by intersecting maximally non-implying subsets, aka. remainders. That is, to remove an element a from a theory A one considers subsets of A that are such that they do not entail a whereas all proper supersets do. This has turned out to be a sustainable mathematical idiom, and it is an idea that will be taken over in this paper. We therefore need to generalise the notion of a remainder to input/output logic. This is entirely straightforward:

Definition 2 *(Remainders sets)* $out(G) \perp (a, b)$ is the set of H such that

1. $H \subseteq out(G)$
2. $(a, b) \notin out(H)$, and
3. If $H \subset I \subseteq G$ then $(a, b) \in I$.

Note that remainders are defined as subsets of the closed set $out(G)$ rather than of the base G, so we are aiming for a version of *theory contraction* rather than, in the terminology of (Hansson 1993), *base-contraction*. An important property of remainders, as so defined, is that they are closed under *out*:

Lemma 7 *If $H \in out(G) \perp (a, b)$ then $H = out(H)$.*

Proof Suppose that $H \in out(G) \perp (a, b)$. We need to show that $H = out(H)$. By inclusion for *out*, it suffices to show that $(c, d) \in H$ whenever $(c, d) \in out(H)$. Suppose therefore that $(c, d) \in out(H)$. By monotony for *out* we have that $out(H) \subseteq out(G)$, hence $(c, d) \in out(G)$. Now, suppose for reductio ad absurdum that $(c, d) \notin H$. Then since $H \in out(G) \perp (a, b)$ and $(c, d) \in out(G)$ we know that $(a, b) \in out(H \cup (c, d))$. But since $(c, d) \in out(H)$ by assumption, we have that $out(H) = out(H \cup (c, d))$ so that $(a, b) \in out(H)$, contrary to hypothesis. □

Input/output logic also enjoys a compactness property that ensures that a non-implying subset of $out(G)$ (w.r.t. to some norm (a, b)) can always be expanded to a *maximally* non-implying subset of G, that is, to a remainder (a proof can be found in ((Stolpe 2008a), chap. 4)):

Lemma 8 (Maximalisability) *If $(a, b) \notin F \subseteq out(G)$ then there is an $F^+ \supseteq F$* such *$F^+ \in out(G) \perp (a, b)$.*

Finally, it is possible to establish a principled relation between different sets of remainders: If F is a subset of $out(G)$ that maximally does not imply (a, b), and F happens not to imply (a, c), then it is also maximally non-implying w.r.t (a, c). To prove this, we shall need to appeal to the following property, which mirrors (Alchourron et al. 1985, lemma 2.1):

Lemma 9 *Let $H := \bigcap(out(G) \perp (a, b))$. Then $out(G, a) \subseteq out(H \cup (a, b), a)$.*

Proof Suppose that $d \in out(G, a)$. We want to show that $d \in out(H \cup (a, b), a)$. By *AND* it suffices to show that $b \to d \in out(H, a)$. Note that, $b \to d \in out(G, a)$, by *WO*, since $d \in out(G, a)$. Suppose for reductio ad absurdum that $b \to d \notin out(F, a)$

for some $F \in out(G) \perp (a, b)$. Then, by the maximality of F it follows that $b \in out(F \cup (a, b \rightarrow d), a)$, so lemma 4 and 7 yield $(a, (b \rightarrow d) \rightarrow b) \in F$. Now,

$$\begin{aligned}(b \rightarrow d) \rightarrow b &= \neg(b \rightarrow d) \vee b \\ &= \neg(\neg b \vee d) \vee b \\ &= (b \wedge \neg d) \vee b \\ &= b\end{aligned}$$

Hence $(a, b) \in F$, by one application of *WO*, contradicting $F \in out(G) \perp (a, b)$. □

This suffices to establish the above-mentioned relationship:

Lemma 10 *If* $F \in out(G) \perp (a, b)$ *then* $F \in out(G) \perp (a, c)$ *for all* $(a, c) \in out(G)$ *with* $(a, c) \notin F$.

Proof Suppose $F \in out(G) \perp (a, b)$ and $(a, c) \notin F$ for $(a, c) \in out(G)$. It will suffice to show that whenever $F \subset F' \subseteq out(G)$ then $(a, c) \in F$. Let $F \subset F' \subseteq out(G)$. Since $F \in out(G) \perp (a, b)$ we have $(a, b) \in F$. But also, since $F \in out(G) \perp (a, b)$ then $out(G) \perp (a, b)$ is non-empty so $\bigcap(out(G) \perp (a, b)) \subseteq F'$. Hence $out(G, a) \subseteq out(F' \cup (a, b), a)$ whence since $(a, c) \in out(G)$ we have $(a, c) \in F$ as desired. □

4 Derogation

Although theory revision was first put on the agenda in the work of the philosophers Harper (1976a, b) and Levi (1977, 1991), Alchourrón and Makinson (1981) are usually given credit for turning it into a well-defined research programme (Williams and Rott 2001).

The basic idea underlying the AGM paradigm, is to break down the revision of a theory into two steps: First the theory is downsized, or *contracted*, to just below the threshold of making it inconsistent with the new data, and then the reduced theory is *expanded* by adding the new proposition and taking the closure of the result. This is conceptually very simple, but from the technical side of things there are certain difficulties to overcome, the principal one being the following: In general there is more than one way to reduce a theory so as to make it consistent with a given input. In most cases the result is not unique, and we obtain a number remainders of the theory to be revised that are *all* compatible with the proposition to be added. Furthermore, neither each subset taken alone (called *maxichoice* contraction), nor their common part (*full meet* contraction) can serve as a reasonable solution; the former preserves too much, whereas the second retains too little information (Alchourron and Makinson 1982; Bochman 2001, p. 323) . The solution devised by Alchourrón et al. is well known; they use a preference mechanism to select only a *subset*—intuitively the preferred ones—among the remainders. Unlike maxichoice and full-meet contraction, the set of remainders scheduled for intersection is *selected*, which adds flexibility to the model. The resulting framework is known as *partial meet* contraction—the paradigm of revision theory.

Adapting this approach to input/output logic requires us first of all to be explicit about what a selection function is:

Definition 3 A selection function for a set of norms G, is any function of type $2^{2^{out(G)}} \mapsto 2^{2^{out(G)}}$, such that

(a) $\emptyset \subset \delta(X) \subseteq X$ if X is a non-empty subset of $2^{out(G)}$, and
(b) $\delta(X) = \{out(G)\}$ otherwise.

A partial meet *derogation* operator can now be defined in the usual way as follows:

Definition 4 $out(G) - (a,b) = \bigcap \delta(out(G) \perp (a,b))$, where δ is a selection function for G.

Intuitively the operation δ chooses the most preferred elements among the set of remainders of $out(G)$ (w.r.t. the derogandum (a, b)), if there are any. If there are none, δ returns the set consisting only of $out(G)$. Note that the selection function δ is entirely abstract. That is, we make no assumptions about *how* δ selects the remainders. It may of course be fleshed out at some later point. Various explicitly defined preference relations from the literature (see e.g. (Alchourron and Makinson 1981; Prakken 1997; Prakken and Sartor 1997; Nebel 1992)) could be used—some of them are specifically tailored for legal reasoning—but it makes sense to leave all options open at the outset.

Most of the properties that are *characteristic* of partial meet *derogation*, as so defined, resemble closely the properties that characterise classical contraction—but there are some interesting new-comers as well:

Theorem 5 *Partial meet derogation satisfies the following properties:*

***D* – *Closure* :** $out(G) - (a, b) = out(out(G) - (a, b))$
***D* – *Inclusion* :** $out(G) - (a,b) \subseteq out(G)$
***D* – *Success* :** If $\nvdash b$ then $(a,b) \notin out(G) - (a,b)$
***D* – *Extensionality* :** If $out((a, b)) = out((c, d))$. then $out(G) - (a, b) = out(G) - (c, d)$
***Input Dependence* :** If $(c,d) \in out(G) \setminus out(G) - (a,b)$ then $a \vdash c$
***Local Recovery* :** If $(c, d) \in out(G)$ and $a \vdash c$ then $(c,d) \in out((out(G) - (a,b)) \cup (c,b))$

Proof We prove only the last two postulates. The remaining proofs can be found in (Stolpe 2008a). The proof of Local Recovery is similar to that of lemma 9 with minor modifications: Suppose that $(c, d) \in out(G)$ and that $a \vdash c$. We want to show that $(c,d) \in out((out(G) - (a,b)) \cup (c,b))$. By *AND* it suffices to show that $(c, b \to d) \in out(G) - (a, b)$. Note that, $(c, b \to d) \in out(G)$, by *WO*, since $(c, d) \in out(G)$. Suppose for reductio ad absurdum that $(c, b \to d) \notin H$ for some $H \in out(G) \perp (a,b)$. Then, by the maximality of H it follows that $(a,b) \in out(H \cup (c, b \to d))$, so lemma 4 and 7 yield $(c \wedge a, (b \to d) \to b) \in H$. Now,

$$\begin{aligned}(b \to d) \to b &= \neg(b \to d) \vee b\\ &= \neg(\neg b \vee d) \vee b\\ &= (b \wedge \neg d) \vee b\\ &= b\end{aligned}$$

Hence $(c \wedge a, b) \in H$, by one application of *WO*. Since $a \vdash c$ it follows that $(a, b) \in H$, contradicting $H \in out(G) \perp (a, b)$. For Input Dependence, suppose $(c, d) \in out(G) \setminus out(G) - (a, b)$. We need to show that $a \vdash c$. It follows from the supposition that there is an $F \in out(G) \perp (a, b)$ with $(c, d) \notin F$. By $F \in out(G) \perp (a, b)$ we have $(a, b) \in out(F \cup (c, d))$ whence $a \vdash c$ by lemma 6 as desired. □

The interpretation of these conditions is mostly straightforward. A-Closure says that derogation of a norm from a system produces a new *system*, that is, derogation produces a set of norms that is closed under the rules that characterise the *out*-operator. D-Inclusion precludes the addition of new norms in the process of derogation, D-Success ensures that the derogandum is no longer in the system after the derogation has been performed (unless the derogandum is a norm with a tautologous consequent), whilst D-Extensionality says that derogating equivalent norms (modulo *out*) from the same code produces the same result—i.e. only the logical content of a norm, not its particular formulation, affects derogation. The postulate of Input Dependence is notable for the fact that it lacks a counterpart in classical AGM-revision theory. It can therefore be regarded as a genuine idiosyncrasy of normative reasoning. Together with Local Recovery, to which I shall return shortly, it expresses the *locality* of normative reasoning. By the locality of normative reasoning I shall mean the fact that only weaker contexts than the one described by the antecedent of the norm to be removed need be taken into consideration in order to remove that norm. Incisions are confined, so to speak, to the context described by the antecedent of the derogating norm. Local Recovery expresses a different aspect of essentially the same phenomenon. It is a restricted version of the recovery postulate from classical revision theory which captures a *global* notion of the same property:

AGM-Recovery: $A \subseteq Cn((A - a) \cup a)$ whenever A is a theory.

This is often referred to as the principle of minimal-mutilation, since it prevents unmotivated retraction of elements. Indeed AGM-Recovery keeps incisions into a theory so small that they can be undone by simply adding the removed sentence (Fuhrmann 1997, p. 43). The intuitive plausibility of this property is controversial, though (see e. g. Makinson 1987; Hansson 1991; Fuhrmann 1991; Levi 1991). Hansson (1999, p. 73) gives the following counterexample (in which theories are interpreted as belief sets):

> I believe that 'Cleopatra had a son' (ϕ) and that 'Cleopatra had a daughter' (ψ), and thus also that 'Cleopatra had a child' ($\phi \vee \psi$, briefly κ). Then, I receive information that makes me give up my belief in κ, and contract my belief set accordingly, forming $A - \kappa$. Soon afterwards I learn from a reliable source that Cleopatra had a child. It seems perfectly reasonable for me to then add κ (i.e. $\phi \vee \psi$) to my set of beliefs without also reintroducing either ϕ or ψ.

It is interesting to note, therefore, that the operation of *derogation*, does not satisfy a corresponding global version of recovery. That is, the following principle is *not* valid:
Global Recovery: $out(G) \subseteq out((out(G) - (a,b)) \cup (a,b))$.

Example 2 (*Counterexample to Global Recovery*) Put $G := \{(a,b),(a \wedge c,b)\}$. Suppose that a and c are logically distinct and that $\nvdash b$. From the latter assumption, it follows by D-Success and *SI* that $(a,b) \notin out(G) - (a \wedge c, b)$. Now, suppose for reductio that $(a,b) \in out((out(G) - (a \wedge c,b)) \cup (a \wedge c,b))$. Then it follows from lemma 6 that $a \vdash c$ contradicting the assumption that a and c are logically distinct.[3]

Local Recovery and Input Dependence are thus aspects of the same phenomenon in the following sense: Whereas *removals* of norms from a system only require incisions into contexts that are *weaker* than the one described by the antecedent of the derogating norm, *additions* of new norms to a system conversely only affects *stronger* contexts. Hence if the norm (a, b) is removed from a system $out(G)$, and (c, b) is added back in, where c is logically weaker than a, then $out(G)$ recovers in all contexts from c up to and including a, but in none weaker than c.

In philosophical terms, the locality of normative reasoning expressed by Input Dependence and Local Recovery, is a consequence of the anti-naturalism w.r.t. norms that is built into the input/output idiom: A norm is simply an agreed upon standard of correctness. Such standards come to exist and pass out of existence in the course of the history of a normative system—sometimes as a consequence of acts of legislation, sometimes as commandments of a trusted authority or leader, sometimes as a result of gradually formed societal habits, customs and traditions, explicit agreement or design (Ross 1968)—they are not platonic universals or logical truths. Stated differently, norms are posited by human acts of will, and, in the words of Kelsen; ‘Norms posited by human acts of will are arbitrary in the genuine signification of the word: that is, they can decree any behaviour whatsoever to be obligatory’ (Kelsen 1991, p. 4).

As mentioned earlier, this view is basic to the input/output idiom where the notion of an arbitrary stipulation is reflected by ordered pairs. Norms are thus not construed as propositions so there is no logical relationship between the norms as such. Compare the counterexample above with a representation in propositional logic: Put $K := \{a \to b, a \wedge c \to b\}$. Then $Cn(K)$ contains $(a \wedge c \to b) \to (a \to b)$ since $(a \to b) \to ((a \wedge c \to b) \to (a \to b))$ is a tautology. A standard result from classical revision theory tells us that $(a \wedge c \to b) \to (a \to b)$ belongs to every subset of $Cn(K)$ which is maximally such that it does not contain $a{\to}b$. This property generalises to all propositions in $Cn(K)$ for arbitrary K, whence classical contraction recovers *globally*: $Cn(K) \subseteq Cn((Cn(K) - (a \wedge c \to b)) \cup (a \wedge c \to b))$. In contrast no statement like $(a,b) \to ((a \wedge c,b) \to (a,b))$ is derivable in input/output logic—indeed such a statement is not even well-formed. But we have $b{\to}(c{\to}b)$ for all *consequents* of norms b, whence input/output logic recovers locally.

[3] The failure of the global version of recovery for contraction of input/output systems was noted in Boella et al. (2009). The authors do not offer an alternative, however.

Note that we could have expressed locality with a single postulate:

Locality: If $(c, d) \in out(G) \setminus out(G) - (a, b)$ then $(c, b \rightarrow d) \in out(G) - (a, b)$.

since this is equivalent to the conjunction of Local Recovery and Input Dependence:

Theorem 6 *Locality is equivalent to the conjunction of Local Recovery and Input Dependence.*

Proof To show that Locality implies Local Recovery, suppose $(c, d) \in out(G)$ and $a \vdash c$. The limiting case where $(c, d) \in out(G) - (a, b)$ is trivial, so suppose $(c, d) \notin out(G) - (a, b)$. Then $(c, b \rightarrow d) \in out(G) - (a, b)$ by Locality. Hence $(c, d) \in out((out(G) - (a, b)) \cup (c, b))$ by *AND*. For the converse direction, suppose $(c, d) \in out(G) \setminus out(G) - (a, b)$. Then by Input Dependence we have $a \vdash c$. Hence, by Local Recovery we have $(c, d) \in out((out(G) - (a, b)) \cup (c, b))$, whence $(c, b \rightarrow d) \in out(G - (a, b))$. □

Hence, we could have obtained a more compact representation of derogation substituting Locality for Input Dependence and Local recovery. There is in fact more than one way to reduce the number of postulates, for we could also have adapted the generalisation from classical revision theory that substitutes *relevance* for recovery:

D-Relevance: If $(c, d) \in out(G) \setminus out(G) - (a, b)$ then there is an F s. t.

1. $out(G) - (a, b) \subseteq F \subseteq out(G)$
2. $(a, b) \notin out(F)$, and
3. $(a, b) \in out(F \cup (c, d))$.

The postulate of relevance (of which D-Relevance is a straightforward translation) was first introduced in classical revision theory by Hansson in (1991). Hansson proves the equivalence of relevance and recovery for contraction on *closed* sets of sentences, and an analogous property holds for derogations:

Theorem 7 *D-Relevance is equivalent to Input Dependence and Local Recovery in the presence of D-Closure, D-Success, D-Inclusion and D-Extensionality.*

Proof It was proved in Stolpe (2008a) that partial meet derogation, as defined in definition 4 is completely characterised by D-Closure, D-Success, D-Inclusion, D-Extensionality and D-Relevance. The equivalence therefore follows as a corollary to the representation theorem 8 below. □

Notwithstanding this equivalence, D-Relevance does not, in my opinion, give a very informative characterisation of the derogation operation (and the same goes for relevance in classical revision theory), since it comes very close to simply stating the condition for being a preferred remainder. Preferred remainders belong to the *semantics*, or the *construction*, of partial meet derogation. Hence the possibility of characterising that construction in terms of D-Relevance is not very surprising. What we would want, ideally, is a representation expressed as far as possible in terms of the membership or non-membership of norms as a consequence of derogation, without appealing directly to auxiliary constructions such as maximally non-implying sets and so forth. Local Recovery and Input Dependence live up to that requirement to a considerable extent. Moreover the separation of concerns between them yields a clearer picture of the idiosyncrasy of normative reasoning than does the more compressed postulate of Locality—hence, the axiomatisation above.[4]

The main result of the present section is the representation theorem that shows the above-mentioned axiomatisation to give a complete characterisation of derogation. In order to prove it, it will be convenient to appeal to postulates that are not in fact in the list themselves. These are:

D-Failure: If $\vdash b$ then $out(G) = out(G) - (a, b)$

and

D-Vacuity: If $(a,b) \notin out(G)$ then $out(G) = out(G) - (a, b)$.

As the next pair of lemmas show, these postulates are implied by other members in the list, taken jointly, so there is no need to expand the set of postulates:

Lemma 11 *D-Closure, D-Inclusion, Input Dependence and Local Recovery imply D-Failure.*

Proof We have $out(G) - (a,b) \subseteq out(G)$ by D-Inclusion. For the other direction suppose suppose $\vdash b$ and $(c, d) \in out(G)$. We want to show that $(c, d) \in out(G) - (a, b)$. Suppose not. Then, by Input Dependence we have $a \vdash c$ so $(c,d) \in out(out(G) - (a,b)) \cup (c,b))$ by Local Recovery. By D-Closure, $\vdash b$ and *SI* therefore, we have $(c, b) \in out(G) - (a, b)$. Hence $out(out(G) - (a,b)) \cup (c,b))$ $= out(G) - (a,b)$, by one more application of D-Closure, so $(c, d) \in out(G) - (a, b)$ as desired. □

Lemma 12 *D-Inclusion, D-Closure, Input Dependence and Local Recovery imply D-Vacuity.*

Proof By D-Inclusion it suffices to show that $out(G) \subseteq out(G) - (a,b)$ on the assumption that $(a,b) \notin out(G)$. So suppose $(c,d) \notin out(G) - (a,b)$. We need to show that $(c,d) \notin out(G)$. Assume the opposite. Then, by Input Dependence we have $a \vdash c$. Thus, since $(c, d) \in out(G)$ and $a \vdash c$ we have $(c,d) \in out((out(G) - (a,b)) \cup (c,b))$ by Local Recovery. Moreover if $(a,b) \notin out(G)$ it follows that $(c,b) \notin out(G)$, by *SI*, whence $(c,b) \notin out(G) - (a,b)$ by D-Inclusion. In other words $out((out(G) - (a,b)) \cup (c,b)) = out(out(G) - (a,b))$ so $out(out(G) - (a, b)) = out(G) - (a, b)$ by D-Closure. Thus, $(c,d) \in out((out(G) - (a,b)) \cup (c,b))$ entails that $(c, d) \in out(G) - (a, b)$ contrary to assumption. □

Necessary is also the following conditional converse of extensionality:

Lemma 13 *If $out(G) \perp (a,b) = out(G) \perp (c,d)$ then $out((a, b)) = out((c, d))$, whenever $(a, b), (c, d) \in out(G)$.*

Proof Suppose for reduction that $out(G) \perp (a,b) = out(G) \perp (c,d)$ while $out((a,b)) \neq out((c,d))$, say $out((a,b)) \nsubseteq out((c,d))$. By the monotony and

[4] Having said that, I shall resort to the use of D-Relevance in the characterisation of revision of sets of norms in Sect. 6, since a better representation of the revision operation has not been forthcoming.

idempotence for *out* we have $(a,b) \notin out((c,d))$. Hence, since $(c, d) \in out(G)$ it follows that $out((c, d))$ can be expanded to a maximal subset F of G such that $(a,b) \notin F$. Thus $F \in out(G) \perp (a,b)$ while $F \notin out(G) \perp (c,d)$ since $(c, d) \in F$, contrary to assumption. □

This suffices to prove that partial meet derogation is characterised by the listed postulates:

Theorem 8 *Every operation that satisfies D-Closure, D-Inclusion, D-Extensionality, D-Success, Input Dependence and Local Recovery is a partial meet derogation operation.*

Proof In the appendix. □

4.1 On the relation between contraction and derogation

Before we turn to the operation of amendment, we should pause to record a few facts about the relationship between derogation, as here defined, and classical contraction, as this relationship may be of some theoretical interest. Theorem 2 shows that input/output logic generalises classical logic, so it is natural to expect a structured relationship. Knowing the precise nature of this relationship, could be very useful for carrying results and techniques over from classical revision theory to the realm of normative reasoning. The results in this section were presented in (Stolpe 2010). They are repeated here, for the sake of expository completeness:

Define a function μ as follows:

Definition 5 Let $\mu : out(G) \perp (a,b) \mapsto 2^L$ be a function such that $\mu(F) = F(Cn(a))$.

Then μ maps remainders of a normative system to remainders of its output under the input of the derogandum:

Lemma 14 If $F \in G \perp (a,b)$ *then* $\mu(F) \in G(Cn(a)) \perp b$.

Proof Suppose $F \in G \perp (a,b)$ and suppose for reduction that $\mu(F) \notin G(Cn(a)) \perp b$. By definition $\mu(F) = F(Cn(a))$, and since $F \in G\perp(a, b)$ we have $F \subseteq G$ whence $F(Cn(a)) \subseteq G(Cn(a))$. There are thus two cases to consider:

1. $F(Cn(a)) \vdash b$: By compactness for logical consequence, there is thus a finite set of rules $(a_1, b_1), \ldots, (a_n, b_n) \in F$ such that $a_i \in Cn(a)$ for each $i \leq n$, and $\bigwedge_{i=1}^{n} b_i \vdash b$. Hence $(a, b) \in out(F)$, by repeated applications of *SI*, *AND* and *WO*, contradicting $F \in G \perp (a,b)$.
2. There is a $B \in G(Cn(a)) \perp b$ such that $F(Cn(a)) \subset B$: It follows that there is a $d \in B \setminus F(Cn(a))$. Since $B \subseteq G(Cn(a))$ we have $(c, d) \in G$ for some $c \in Cn(a)$. Clearly $(c,d) \notin F$ so $(a,b) \in out(F \cup (c,d))$ by the membership of F in $G \perp (a,b)$. Now, since $b \notin B \supset F(Cn(a))$ it follows that $(a,b) \notin F$ whence $(a,b) \in out(F \cup (c,d)) \setminus out(F)$. By lemma 6 we therefore have $a \vdash c$, and by lemma 4 we have $(a, d \rightarrow b) \in out(F)$. Hence $F(Cn(a)) \vdash d \rightarrow b$ so $B \vdash d \rightarrow b$ by monotony for classical logic. Since $d \in B$ therefore, it follows that $B \vdash b$, contrary to the assumption that $B \in G(Cn(a)) \perp b$. □

This mapping is onto:

Lemma 15 *μ is surjective.*

Proof We want to show that $B = \mu(F)$ for every $B \in G(Cn(a)) \perp b$ and some $F \in G \perp (a,b)$. So, suppose that $B \in G(Cn(a)) \perp b$ and put $F := \{(c,d) \in G : B \vdash d$ and $c \in Cn(a)\}$. We first show that $B = F(Cn(a))$. For the left-in-right inclusion suppose $d \in B \subseteq G(Cn(a))$ then $(c, d) \in G$ for some $c \in Cn(a)$, so $(c, d) \in F$ by the construction of F. It follows that $d \in F(Cn(a))$. The converse inclusion is immediate from the construction. Next we show that $F \in G \perp (a,b)$. Since $b \notin B$ by the assumption that $B \in G(Cn(a)) \perp b$, it follows that $(a,b) \notin F$ since $B = F(Cn(a))$. Moreover, $F \subseteq G$ by the construction of F so it may be expanded to an $F' \subseteq G$ such that $F' \in G \perp (a,b)$, by lemma 8. We show that $F(Cn(a)) = F'(Cn(a))$. The left in right inclusion is immediate. For the converse inclusion, suppose $F'(Cn(a) \not\subseteq F(Cn(a))$. Then there is a $d \in F'(Cn(a)) \setminus F(Cn(a))$, whence $d \notin B = F(Cn(a))$. It follows that $B \cup d \vdash b$ so $F(Cn(a)) \cup d \vdash b$, whence $F(Cn(a)) \vdash d \rightarrow b$. In other words, we have $d \rightarrow b \in Cn(F(Cn(a)))$, which means that $(a, d \rightarrow b) \in out(F) \subseteq out(F')$. But then $(a, b) \in out(F')$ by *AND*, since $d \in F'(Cn(a))$ by assumption, contradicting $F' \in G \perp (a,b)$. □

and one-to-one:

Lemma 16 *μ is injective.*

Proof We need to show that $\mu(F) = \mu(F')$ implies $F = F'$ for $F, F' \in G \perp (a,b)$. Suppose for reductio that $F(Cn(a)) = F'(Cn(a))$, but $F \neq F'$. Assume without loss of generality, that $(c,d) \in (F' \setminus F) \neq \emptyset$. Then $(a,b) \in out(F \cup (c,d))$ so $a \vdash c$ and $(a \wedge c, d \rightarrow b) \in F$, whence $(a, d \rightarrow b) \in out(F)$. Moreover, $(a, d) \in out(F')$, by *SI*, since $(c, d) \in F'$. It follows that we have $F(Cn(a)) \cup F'(Cn(a)) \vdash d$ and $F(Cn(a)) \cup F'(Cn(a)) \vdash d \rightarrow b$. However, since $F(cn(a) = F'(Cn(a))$ this is tantamount to saying $F'(Cn(a)) \vdash d$ and $F'(Cn(a)) \vdash d \rightarrow b$. Therefore $F'(Cn(a)) \vdash b$, whence $(a, b) \in out(F')$, contradicting $F' \in G \perp (a,b)$. □

Thus, we have the following corollary:

Corollary 1 $out(G) \perp (a,b)$ and $out(G,a) \perp b$ are in one-to-one correspondence.

Hence, every derogation operation can be represented by a contraction operation in the following sense:

Theorem 9 *For every partial meet derogation operator $-$ there is a partial meet-contraction operation $\backsim$ such that $out(out(G) - (a,b), a) = out(G,a) \backsim b$.*

Proof Suppose $out(G) - (a,b) = \bigcap \delta(out(G) \perp (a,b))$. Choose $\gamma(out(G,a) \perp b)$ to be the set $\{B \in out(G,a) \perp b : \mu^{-1}(B) \in \delta(out(G) \perp (a,b))\}$. It suffices to show that $(a,d) \in \bigcap \delta(out(G) \perp (a,b))$ iff $d \in \bigcap \gamma(out(G,a) \perp b)$. For the left-to-right direction suppose $d \notin \bigcap \gamma(out(G,a) \perp b)$. Then there is a $B \in \gamma(out(G,a) \perp b)$ such that $d \notin B$. By the definition of γ we have that $\mu^{-1}(B) \in \delta(out(G) \perp (a,b))$. Now since $d \notin B$ it follows that $(a,d) \notin \mu^{-1}(B)$ so $(a,d) \notin \bigcap \delta(out(G) \perp (a,b))$ as desired. For the converse direction, suppose $(a,d) \notin \bigcap \delta(out(G) \perp (a,b))$.

Then there is an $F \in out(G) \perp (a,b)$ such that $(a,d) \notin F$. It follows that $d \notin out(F,a) = \mu(F)$. By lemma 14, $\mu(F) \in out(G,a) \perp b$, whence $\mu(F) \in \gamma(out(G,a) \perp b)$, by the definition of γ. It follows that $d \notin \bigcap \gamma(out(G,a) \perp b)$, which completes the proof. □

and vice versa:

Theorem 10 *For every partial meet contraction operator $\backsim$, there is a partial meet derogationoperator—such that $out(G,a) \backsim b = out(out(G) - (a,b),a)$ there is a partial meet-contraction operation $\backsim$ such that $out(out(G) - (a, b), a) =$.*

Proof Similar in all its essentials to the preceding one. □

Hence,

Corollary 2 An operation $-$ is a partial meet derogation operation iff there is a partial meet contraction operation $\backsim$ such that $out(G,a) \backsim b = out(out(G) - (a,b),a)$.

The function μ thus provides a convenient bridge that can be used to explore aspects of norm-system dynamics that have well-understood counterparts in classical revision theory. See (Stolpe 2010) for an example. Another interesting possibility would be to study the priority relation over $out(G)$ induced by a derogation operation $-$. The counterpart in classical revision theory is known as *epistemic entrenchment* (see e.g. Makinson and Gärdenfors 1988; Rott 1992; Bochman 2001). Epistemic entrenchment is interesting for several reasons. First, it yields a *constructive* definition of partial meet contraction, and secondly it gives insights into the basic properties of preference that are determined by the logical properties alone of the items so related. It goes without saying that a similar analysis applied to norms, would be of great value for understanding the mechanisms of norm-system dynamics that require conflict resolution, such as indeed predicaments, contrary-to-duty situations and permission. I leave this for future research.

5 Amendment

A well-known principle called the Levi-identity in (Gärdenfors 1988) after Isaac Levi, identifies a *revision* of a theory with respect to a proposition a with the contraction of $\neg a$ followed by the addition of a (and the subsequent closure of the result). The basic idea is to understand revision as *consistency preserving* addition, i. e. revision is construed as a two-step process; first remove enough of the original theory to ensure that the outcome is consistent with the scheduled addition, then add the new element and close the result. Generalising this principle to input/output logic suggests the following definition:

Definition 6 $out(G) \dotplus (a,b) := out((out(G) - (a,\neg b)) \cup (a,b))$.

I shall call this operation an operation of *amendment* in order to keep it separate from classical *revision*. The reader should note though that the term 'amendment' in its legal sense usually refers to a change made to a previously adopted or pending bill, or to a change made to a written constitution. Definition 6, in contrast, does not presuppose the existence of any norm, and, of course, may require such norms, if they do exist, to be suspended. As with the operator of derogation, the legal term ‘amendment’ may be regarded as a special case of the corresponding notion defined in this paper, that is, as a special case of norm-system revision.

Note that definition 6 interprets amendment as *locally* consistent expansion. Thus, norm-system revision does not require coherence across contexts generally. This is as it should be. The point of adding a new *norm* (i.e. a *conditional* directive) to a code is to regulate the circumstances expressed by *that* condition of application. Nevertheless, as the postulates of Input Dependence and Local Recovery show, revising a context may require revision of logically weaker contexts, which will in turn affect logically stronger contexts (by the rule of input strengthening). However, logically independent contexts will never be affected. We may see this as an expression of one important idiosyncrasy of normative reasoning; it is localised. The locality of normative reasoning, on the analysis given here, is ultimately due to the fact that norms are construed as arbitrary stipulations—there is no necessary *logical* relationship between antecedents and consequents of norms, they are simply decreed to hold. As argued in Sect. 3.1 the negation of a norm should should therefore be understood as local not global negation, although technically only the latter is a lattice-theoretic complement in the set of all norms.

The principal properties of amendment as defined in 6 are given by the next theorem:

Theorem 11 $\dot{+}$ satisfies all of the following:

1. ***A − Closure :*** $out(G)\dot{+}(a,b) = out(out(G)\dot{+}(a,b))$
2. ***A − Inclusion :*** $out(G)\dot{+}(a,b) \subseteq out(G \cup (a,b))$
3. ***A − Success :*** $(a,b) \in out(G)\dot{+}(a,b)$
4. ***A − Consistency :*** If $(a,f) \in G\dot{+}(a,b)$ then $b \vdash f$
5. ***A − Extensionality :*** If $out((a,\ b)) = out((c,\ d))$ and $a \equiv c$ then $out(G)\dot{+}(a,b) = out(G)\dot{+}(c,d)$.
6. ***A − Relevance :*** If $(c,d) \in out(G) \setminus out(G)\dot{+}(a,b)$, then there is an F such that,
   (a) $(out(G)\dot{+}(a,b)) \cap out(G) \subseteq F \subseteq out(G)$,
   (b) $(a,\neg b) \notin out(F)$, and
   (c) $(a,\neg b) \in out(F \cup (c,d))$.

Proof We prove A-Extensionality and A-Relevance only, all others are entirely obvious: For A-Extensionality, suppose $out((a,\ b)) = out((c,\ d))$ and $a \equiv c$. Then by lemma 2 we have that $b \equiv\ d$. Hence $out(G)\dot{+}(a,b) = out((out(G) - (a,\neg b)) \cup (a,b)) = out((out(G) - (c,\neg d)) \cup (c,d)) = out(G)\dot{+}(c,d)$, by D-Extensionality together with *WO* and *SI*. For A-Relevance, suppose $(c,d) \in out(G) \setminus out(G)\dot{+}(a,b)$. We need to find an F such that;

1. $((out(G)\dot{+}(a,b)) \cap out(G) \subseteq F \subseteq out(G)$.

2. $(a, \neg b) \notin out(F)$, and
3. $(a, \neg b) \in out(F \cup (c, d))$.

Since amendment has been defined by the Levi identity we have $out(G) \dotplus (a, b) = out((G - (a, \neg b)) \cup (a, b))$. By assumption $(c, d) \notin out(G) \dotplus (a, b)$, so $(c, d) \notin out(G) - (a, \neg b)$ by the closure properties of the *out* operator. By D-Relevance and the assumption that $(c, d) \in out(G)$ there is thus an F such that

(a) $out(G) - (a, \neg b) \subseteq F \subseteq out(G)$,
(b) $(a, \neg b) \notin out(F)$, and
(c) $(a, \neg b) \in out(F \cup (c, d))$.

It follows from (a) and maximalisability (lemma 8) that F can be extended to an $F' \in out(G) \perp (a, \neg b)$. By the definition of remainders we thus have that $(a, \neg b)$ / $\in F' = out(F')$ by lemma 7, so condition 2 holds. Morover $(a, \neg b) \in out(F \cup (c, d)) \subseteq out(F' \cup (c, d))$ since $F \subseteq F'$, so condition 3 holds too. It remains to show condition 1: Note first that $(a, b) \in F'$ if $(a, b) \in out(G)$. Suppose not. Then, since $F \in out(G) \perp (a, \neg b)$ it follows that $(a, \neg b) \in out(F' \cup (a, b))$. Hence $(a, b \rightarrow \neg b) = (a, \neg b) \in F'$, contrary to condition 2. As a result $out(G) \cap (a, b) \subseteq F'$. By (a) we therefore have $(out(G) - (a, \neg b)) \cup (out(G) \cap (a, b)) \subseteq F'$. Now,

$$
\begin{aligned}
&(out(G) - (a, \neg b)) \cup (out(G) \cap (a, b)) \\
&\quad = ((out(G) - (a, \neg b)) \cup out(G)) \cap ((out(G) - (a, \neg b)) \cup (a, b)) \\
&\qquad \text{by distribution for } \cup \\
&\quad = out(G) \cap ((out(G) - (a, \neg b)) \cup (a, b)) \text{ by D-Inclusion} \\
&\quad = out(G) \cap (out(G) \dotplus (a, b)) \text{by the definition of } \dotplus
\end{aligned}
$$

So $out(G) \cap (out(G) \dotplus (a, b)) \subseteq out(F') \subseteq out(G)$, as desired. □

The interpretation of these properties is again fairly straightforward: A-Closure says that amendment produces a new normative system, that is, it produces a set of norms that contains all the norms it entails. According to A-Inclusion amendment never transcends simple expansion. This can be seen as a simple relevance criterion; no material that is not implied by the code and the new norm will be included in the result of amending the code with that norm. According to A-Success amendment is always effective—the norm scheduled for addition will always be included in the result. One way to read this is to say that the new norm always takes priority over norms already in the code. A-Success therefore expresses a simple kind of *lex posterior derogat lex priori* principle. A-Consistency says that amendment always produces a code that is *locally* consistent with respect to the applicability condition of the norm to be added, unless that norm is itself locally inconsistent. A-Extensionality says that amending a code with equivalent norms produces similar results. Note though that this property is subject to a proviso. Extensionality does not hold simpliciter, as the following example shows:

Example 3 Suppose $out((a, b)) = out((c, d))$ and that a is not equivalent to c. Then $\vdash b$ and $\vdash d$. Hence $out(G) \dotplus (a, b) = out((out(G) - (a, \neg b)) \cup (a, t)) =$

$out((out(G)- \quad (a,\neg b)) = out(G) - (a,f)$. Similarly $out(G)\dotplus(c,d) = out(G)-(c,f)$. Let a and b be logically independent and put $G := \{(a,f), (b,f)\}$. Then $(a,f) \in out(G) - (c,f) \setminus out(G) - (a,f)$ so $out(G)\dotplus(a,b) \neq out(G)\dotplus(c,d)$.

The example shows that extensionality without the proviso fails for norms with tautologous consequents. The reason is that the equivalence of such norms is not preserved by local negation. That is, if b and d are both tautologies we can have $out((a, b)) = out((c, d))$ and $out((a,\neg b) \neq out((c,\neg d))$. The significance of this non-preservation property is easy to miss. It is an example of one of the many quirks that makes input/output logic different from classical logic, and an example of one of the more subtle properties of amendment that surfaces when there is a construction to back up the postulates. Such a constructive account is lacking in (Boella et al. 2009) and may be partly responsible for the fact that the authors put forth the unqualified version of extensionality as a candidate principle for norm-system revision. The analysis presented here, in contrast, gives a strong reason not to regard that principle as valid. Turning finally to A-Relevance, this principle says pretty much the same about amendment as D-Relevance does about derogation (recall that derogation, according to theorem 7, may be characterised using D-Relevance instead of Input Entailment and Local Recovery), namely that an element that does not *contribute* to the derivation of the local negation of the norm scheduled for addition will never be eliminated. In other words, change is minimal. In fact A-Relevance simply passes the responsibility for keeping changes small on to the underlying derogation operator, for as it turns out $(out(G)\dotplus(a,b)) \cap out(G) = out(G) - (a,\neg b)$ whenever $\dotplus$ is defined by the Levi identity:

Lemma 17 $out(G) \cap out((out(G) - (a,b)) \cup (a,\neg b)) = out(G) - (a,b)$.

Proof The right-in-left inclusion follows immediately from D-inclusion and the closure properties of the *out* operator. For the converse, suppose $(c,d) \notin out(G) - (a,\neg b)$. We need to show that $(c,d) \notin out(G) \cap out((out(G) - (a,\neg b)) \cup(a,b))$. If $(c,d) \notin out(G)$ then we're done. We may therefore suppose that $(c, d) \in out(G)$ whence it suffices to show that $(c,d) \notin out((out(G) - (a,\neg b)) \cup(a,b))$. Suppose for reductio ad absurdum the opposite. Since we have assumed that $(c,d) \in out(G) \setminus out(G) - (a,\neg b)$, D-Relevance tells us that there is an F such that

1. $out(G) - (a,\neg b) \subseteq F \subseteq out(G)$,
2. $(a,\neg b) \notin out(F)$, and
3. $(a,\neg b) \in out(F \cup (c,d))$.

From condition 2 and 3 it follows that $a \vdash c$, by lemma 6. Hence, $out((out(G) - (a,\neg b)) \cup (a,b)) \subseteq out((out(G) - (a,\neg b)) \cup (c,b))$, by *SI*, so $(c, b \rightarrow d) \in out(G) - (a,\neg b)$ by lemma 4, since $(c,d) \in out((out(G) - (a,\neg b)) \cup (c,b))$ by assumption. It follows by condition 1 that $(c, b\rightarrow d) \in F$. Hence, it suffices by *AND* to show that $(c,\neg b\rightarrow d)\in F$, because then $(c, d) \in out(F)$ so $out(F) = out(F \cup (c,d))$ whence condition 2 contradicts condition 3. Suppose therefore that $(c,\neg b \rightarrow d) \notin F$. Then $(c,\neg b \rightarrow d) \notin out(G) - (a,\neg b)$ by condition 1. Since we have $(c,\neg b \rightarrow d) \in out(G)$ from $(c, d) \in out(G)$ by *WO*, there is thus an $H \in$

$\delta(out(G) \perp (a, \neg b))$ such that $(c, \neg b \to d) \notin H$. It follows that $(a, \neg b) \in out(H \cup (c, \neg b \to d))$. Hence $(a \wedge c, (\neg b \to d) \to \neg b) \in out(H)$ by lemma 4. Since $a \vdash c$ it follows that $(a, (\neg b \to d) \to \neg b) \in out(H)$. Now $(\neg b \to d) \to \neg b = \neg b$ so $(a, \neg b) \in H$ by lemma 7 contradicting $H \in out(G) \perp (a, \neg b)$. □

The identity recorded in lemma 17 is a special case of the principle known as the Harper identity from classical revision theory. This principle was named after William Harper who proposed that just as revision is definable in terms of contraction by the Levi identity, contraction may conversely be definable in terms of revision as that which is common to the result of the revision and the original set. Lemma 17 shows that the Harper identity holds, as a derived property, between any operator of amendment $\dotplus$ and its underlying derogation operator. More specifically, when $\dotplus$ is defined by the derogation operation—via the Levi Identity, the Harper identity does indeed produce a contraction operation, and that operation is just the derogation operation we started with. This is a standard result of classical revision theory (cf. Gärdenfors and Rott 1995, p. 57). Its significance is well known, and I shall return to it shortly. Suffice it for now to note that a similar argument can be made w.r.t. to the Levi identity when we start with an arbitrarily chosen amendment operation:

Lemma 18 *Let $\dotplus$ be any operation that satisfies the postulates from theorem 11. Then we have $out((out(G) \cap (out(G) \dotplus (a, b))) \cup (a, b)) = out(G) \dotplus (a, b)$.*

Proof For the left-in-right inclusion it suffices, by monotony and idempotence for *out* together with A-Closure, to show that $(out(G) \cap (out(G) \dotplus (a, b))) \cup (a, b) \subseteq out(G) \dotplus (a, b)$. But $(out(G) \cap (out(G) \dotplus (a, b))) \subseteq out(G) \dotplus (a, b)$, by general set-theory, and $(a, b) \in out(G) \dotplus (a, b)$, by A-Success, so this is immediate. We prove the converse by contraposition and reductio ad absurdum: Suppose $(c, d) \notin out((out(G) \cap (out(G) \dotplus (a, b))) \cup (a, b))$, whilst $(c, d) \in out(G) \dotplus (a, b)$. Then it follows that $(c, d) \notin out(G)$. By A-Inclusion $out(G) \dotplus (a, b) \subseteq out(G \cup (a, b))$, so $(c, d) \in out(G \cup (a, b)) \setminus out(G)$. Consequently there is a subset F of $out(G)$ such that $(c, d) \notin out(F)$ and $(c, d) \in out(F \cup (a, b))$. Suppose that $F \nsubseteq out(G) \dotplus (a, b)$ for every such F. Then since $out(G) \dotplus (a, b) \subseteq out(G \cup (a, b))$ we have $(c, d) \notin out(G) \dotplus (a, b)$ contrary to our second assumption. There is an F, therefore, such that $F \subseteq out(G) \dotplus (a, b)$ and $F \subseteq out(G)$ with $(c, d) \in out(F \cup (a, b))$. By monotony for *out* it follows that $(c, d) \in out((out(G) \cap (out(G) \dotplus (a, b))) \cup (a, b))$, contrary to assumption. □

We are now in position to reproduce the central representation theorem for classical revision: If an amendment operation satisfies the properties listed in theorem 11, then it can be constructed from some partial meet derogation operator via the Levi identity:

Theorem 12 *If an amendment operation $\dotplus$ satisfies A-Closure, A-Inclusion, A-Success, A-Consistency, A-Extensionality and A-Relevance, then it is definable by some derogation operation via the Levi identity.*

Proof In the appendix. □

Taking stock we have the following situation: Theorem 5 shows that the Levi identity maps the set of partial meet derogation operators *into* the set of operators that satisfy the amendment postulates. Theorem 12 shows that this relation is not merely into but *onto*: Every operator that satisfies the amendment postulates can be constructed from *some* partial meet derogation operator by the Levi identity. Lemma 17 next shows that the Harper identity—according to which the derogation of a norm (a, b) from a code is identified with the intersection of that code with its *revision* with $(a, \neg b)$—provides a way back; what the former does the latter undoes. Moreover, applying theorem 12 once more, it follows that the Harper identity is onto too. Hence, derogation and amendment operators are in one-to-one correspondence with the Harper and Levi identities as inverse bijective maps. This can all be set out in a more compressed idiom as follows:

Theorem 13 *Let h and l be operations turning an amendment operation into a derogation operation and vice versa, via the Harper and Levi identities respectively. Then h and l are both onto.*

Proof That l is onto is theorem 12. To show that h is onto suppose $\sim$ is a partial meet derogation operator. We need to find a partial meet amendment operator $+$ such that $h(+) = \sim$. Put $+ := l(\sim)$, then we have the desired result by lemma 17. □

Theorem 14 *Let h and l be defined as in theorem 13. Then h and l are inverses.*

Proof Let $h(l(-)) = \backsim$. Then :

$$\begin{aligned} out(G)\sim(a,b) &= out(G)\cap(out(G)\dotplus(a,\neg b)) && \text{for some } \dotplus \text{ by def. of } h\\ &= out(G)\cap out((out(G)-(a,b))\cup(a,\neg b)) && \text{by the definition of } l\\ &= out(G)-(a,b) && \text{by lemma 17} \end{aligned}$$

so $- = \sim$. Similarly, let $l(h(\dotplus)) = +$. Then:

$$\begin{aligned} out(G)+(a,b) &:= out((out(G)-(a,\neg b))\cup(a,b)) && \text{for some } - \text{ by def. of } l\\ &= out((out(G)\cap(out(G)\dotplus(a,b)))\cup(a,b)) && \text{by the definition of } h\\ &= out(G)\dotplus(a,b) && \text{by lemma 18} \end{aligned}$$

Hence $\dotplus = +$ and we are done. □

What this means is essentially that amendment and derogation are interchangeable idioms. We may view norm-system dynamics in terms of amendment, treating derogation as the derived notion, or vice versa, it doesn't matter. Any method for constructing an operator that satisfies the one list of properties automatically yields a construction for an operator that satisfies the other list. This is a standard result from classical revision theory. It is nice to know that it carries over into input/output logic, and it may be taken as confirmation that the core theory of norm-system dynamics is complete.

6 Sketch of an application: calculating permissions implicit in a code

The aim of the present section is to round off this paper by giving an example of how the theory may be applied. It is hoped that this will suffice to convince the reader of the utility of the general vantage point provided by the theory. This is a point of view according to which norms must be analysed in the larger context of the system to which they belong, rather than, as logicians are wont to, in isolation. Stated differently, it is a salient feature of the theory set out in the preceding pages that there is no logic of norms without attention to the overall behaviour of the system, which includes its modes of transformation. It is my opinion that this gives a more holistic picture—although, to be sure, an idealised one—of the relationship between a code and the norms it contains, and that it can fruitfully be applied to the analysis of concepts in analytical jurisprudence and theoretical computer science alike. As an illustration, the material that follows gives as skeleton theory of what it means for a permission to be implied by a code of norms. The reader should note that this material has been published before (cf. Stolpe 2010a, b). However, since it was precisely the problem of calculating implicit permission that motivated the investigations in the preceding sections to begin with (cf. Stolpe 2008a) a brief recapitulation in the context of the present theory should be warranted.

Now, it is generally agreed that permitted actions fall under one of two broad kinds; those that are *negatively* permitted and those that are *positively* permitted or permitted by decree. Negative permission is simple and denotes an absence of a contrary regulation—in criminal law it is known as the principle *nullum crimen sine lege*. Now, an interesting subset of such actions are those that are *antithetically* permitted, where an antithetically permitted act is understood as one that cannot be prohibited by a code without making that code contradict a positive permission. The exact nature of the latter two concepts, and the relationship between them, is an interesting question to which I shall try to provide an answer.

Let's say, to begin with, that an explicitly permitted action is one that a code explicitly pronounces to be permitted or one that is implicit in what has thus been pronounced. This immediately raises the following problem: What is the nature of this implication relation? Which actions ought to count as implied, and under what circumstances?

Consider the following example from Sect. 8 of the Norwegian personal information act:

Section 8. Personal information may only be processed by the consent of the registered person, or if processing is statutorily warranted, or such processing is required in order to

(a) honour an agreement with the registered person, or to perform a task that accords with the registered person's wishes before such an agreement was entered into,

(b) fulfil a legal obligation on the part of the person responsible for handling the information,

(c) attend to the registered persons vital interests,

(d) perform a task in the interest of the general public,

(e) exercise public authority, or
(f) attend to a justifiable interest that is not outweighed by the regard for the registered person's right to privacy.

This pattern is typical: The statute lays down a general prohibition, and then goes on to list a set of cases for which the prohibition is suspended. In other, words the explicitly permitted actions function as *exceptions* to the general ban.

This kind of positive permission may be called *exemptions*, since what they do is to override a general prohibition in particular elect cases. Several writers have noticed this feature, for instance Ross who says: "norms of permission have the normative function only of indicating, within some system, what are the exceptions from the norms of obligation of the system" (Ross 1968, p. 120). Ross is not, however, correct in claiming that permissive norms necessarily exempt from *already existing* mandatory norms. Consider for instance constitutional rights. A constitutional guarantee does usually not override a prohibition that actually exists. Rather, it is meant to reject *in advance* prohibitions that could conceivably be passed. We may call this class of permissive norms *antithetic permissions* since what they do is to *prevent the code from growing* in such a manner as to prohibit a contrary course of action.

Now, we wish to analyse permissions in the larger context of a system, i.e. to study the interplay between permission and mandatory norms against the background of the sum total of such norms. Hence the proper unit of analysis is a pair $\langle G, P \rangle$ consisting of a set of mandatory norms G and a set of explicitly pronounced permissions P. Both are binary relations over L. I shall say that an item (a, b) is mandatory or permitted (as the case may be) *according to* or *in* such a code, meaning that b is permitted or mandatory whenever a is *true*. It seems natural to analyse the class of permissions I have called exemptions as follows:

Definition 7 (*Exemption*) (a, b) is an exemption according to $\langle G, P \rangle$ iff $(a, \neg b) \in out(G) \setminus out(G) - (a, \neg d)$ for some $(a, d) \in P$.

I shall assume for the remainder of this section that the derogation operation in question is a full meet operation. This is not free of problems though, as the full-meet operation is very strong. However, as is shown in (Stolpe 2010), these problems can be overcome in a way that preserves all the definitions and relationships that are presented here. The account is therefore sound in general outline, and should serve well as an illustration of how norm-system dynamics can be employed in elucidating concepts of relevance to analytical jurisprudence.

Now, definition 7 casts exemptions as cut-backs on the code required to respect the explicit permissions in P. More precisely (a, b) is deemed an exemption if the code contains a prohibition that regulates the state of affairs a by prohibiting b, and $(a, \neg b)$ is such that, unless it is removed, the code will contradict an explicit permission in P. To see how this works, consider the following example.

Example 4 Put $G := \{(t, \neg p)\}$ and $P: = \{(c, p)\}$. Think of these norms as a general prohibition against processing personal information and as an exception for express consent respectively. We have $(c, \neg p) \in out(G)$ by input strengthening. By

D-Success for full meet derogation, however, $(c, \neg p) \notin out(G) - (c, \neg p)$, so (c, p) constitutes an exemption.

Turning now to the concept of antithetic permission, the idea is to see (a, b) as permitted whenever, given the mandatory norms in G, we can't forbid b under the condition a without thereby committing ourselves to forbid, under a condition c that could possibly be fulfilled, something d which is implicit in what has been explicitly permitted (Makinson and van der Torre 2003). Another way to put it is to say that antithetic permissions prevent the set of mandatory norms from growing in such a way as to render explicitly permitted actions forbidden. This checked-growth perspective may be brought out as follows.

Definition 8 (*Antithetic permission*) (a, b) is antithetically permitted according to $\langle G, P \rangle$ iff $(a, \neg b) \notin out(G \cup (a, \neg b)) - (a, d)$ where $(a, d) \in P$.

The following lemma gives an equivalent representation in terms of the norms in $out(G)$ exclusively:

Theorem 15 *(a, b) is antithetically permitted in $\langle G, P \rangle$ iff $(a, \neg b \rightarrow \neg d) \in out(G)$ where $(a, d) \in P$.*

The next example gives a simple illustration of the behaviour of this concept:

Example 5 Put $G := \{(a, d \rightarrow b)\}$ and $P := \{(a, d)\}$. Then (a, b) is antithetically permitted since $(a, \neg b \rightarrow \neg d) \in out(G)$ and $(a, d) \in P$. However, (a, b) is not an exemption, since $(a, b) \notin out(G)$.

As the example shows, antithetic permission does not coincide with exemption, but there is obviously a quite close relationship between them. We are now in position to bring this relationship out clearly, and *a fortiori* to answer (in part) the question we started with:

Theorem 16 If *(a, b) is antithetically permitted in $\langle G, P \rangle$ iff $\langle G \cup (a, \neg b), P \rangle$.*

Proof This is theorem 2 of (Stolpe 2010). □

Thus, the analysis shows that antithetically permitted actions are exemptions in a larger code. Stated differently, antithetic permissions are actions that will be exempted from an operative ban if added to the code. This agrees well with intuition, and also finds support in the sources:

> This is what happens with constitutional rights and guarantees the constitution rejects in advance certain norm-contents (that would affect basic rights), preventing the legislature from promulgating this norm-content, for if the legislature promulgates such a norm-content, it can be declared unconstitutional by the courts and will not be added to the system (Alchourron and Bulygin 1981, p. 397–398).

Theorem 16 thus spells out, with welcome precision, what it means for a permissive provision, such as e.g. a constitutional guarantee, to reject a norm *in advance*, as Alchourrón and Bulygin put it. Many other deductions can be made from this simple analysis, but this will have to do for now. The interested reader should consult (Stolpe 2010).

Acknowledgments This work was partially funded by the Semicolon project supported by the Norwegian Research Council, contract no. 183260.

Appendix

Proof of theorem 8

Proof We need to show that any operation '–' satisfying the listed properties coincides with a partial meet derogation operation. That is, we need to show the existence of a selection function δ such that

$$out(G) - (a,b) = \bigcap \delta(out(G) \perp (a,b))$$

We construct the selection function δ, using a familiar technique, as follows:

$$\delta(out(G) \perp (a,b)) = \begin{cases} \{out(G)\} & \text{if } out(G) \perp (a,b) = \emptyset, \text{ otherwise} \\ \{F \in out(G) \perp (a,b) : out(G) - (a,b) \subseteq F\} \end{cases}$$

We need to show, first of all, that δ is well-defined and that it is a selection function: **Well-definedness**. Starting with well-definedness we need to show that $\delta(out(G) \perp (a,b)) = \delta(out(G) \perp (c,d))$ whenever $out(G) \perp (a,b) = out(G) \perp (c,d)$. So assume that $out(G) \perp (a,b) = out(G) \perp (c,d)$.

In the limiting case that $out(G) \perp (a,b) = \emptyset = out(G) \perp (c,d)$, we have $\delta(out(G) \perp (a,b)) = \{out(G)\} = \delta(out(G) \perp (c,d))$, by the first case of the definition of δ, so we may assume that $out(G) \perp (a,b) \neq \emptyset \neq out(G) \perp (c,d)$.

Proceeding on that assumption, we we turn to the case where either (a, b) or (c, d), say wlog. (a, b), is not in $out(G)$. Then $out(G) \perp (a,b) = \{out(G)\}$, whence $out(G) \perp (c,d) = \{out(G)\}$ as well. Hence, it suffices to show that $\delta(out(G) \perp (g,h)) \neq \emptyset$ for any (g, h) whenever $out(G) \perp (g,h) \neq \emptyset$. The proof splits into two cases:

1. Suppose $(g,h) \notin out(G)$, then $out(G) \subseteq out(G) - (g,h)$ by D-Vacuity. Hence $out(G) \in \delta(out(G) \perp (g,h))$ by the second case of the definition of δ, and we are done.
2. Suppose on the contrary that $(g, h) \in out(G)$. Since we are assuming that $out(G) \perp (g,h)$ is non-empty it follows that $\nvdash h$ by the properties of *out*-entailment. By D-Success, therefore, it follows that $(g,h) \notin out(G) - (g,h)$, whence, by D-inclusion and lemma 8, $out(G) - (g, h)$ can be expanded to a subset F of $out(G)$ such that maximally $(g,h) \notin out(F)$. It follows that $F \in out(G) \perp (g,h)$ and since $out(G) - (g,h) \subseteq F$ we also have $F \in \delta(out(G) \perp (g,h))$. Hence $\delta(out(G) \perp (a,b)) = \delta(out(G) \perp (c,d))$ as desired.

This completes the limiting cases. □

Now, for the principal case where $out(G) \perp (a,b) = out(G) \perp (c,d) \neq \emptyset$ and (a, b), $(c, d) \in out(G)$, note that $F \in \delta(out(G) \perp (a,b))$ implies $out(G) - (a,b) \subseteq F$ by the second case of the definition of δ. Since $out(G) \perp (a,b) = out(G) \perp (c,d)$ and (a, b), $(c, d) \in out(G)$ it follows, by lemma 13 that $out((a, b)) = out((c, d))$, whence $out(G) - (a, b) = out(G) - (c, d)$ by D-Extensionality. Hence $out(G) - (c,d) \subseteq F$, so $F \in \delta(out(G) \perp (c,d))$ by the definition of δ. Therefore $\delta(out(G) \perp (a,b)) \subseteq \delta(out(G) \perp (c,d))$. The other direction is similar so δ is well-defined.

δ is a selection function. To prove that δ is a selection function in the sense of definition 3, we need to show that $\emptyset \subset \delta(out(G) \perp (a,b)) \subseteq out(G) \perp (a,b)$ whenever $out(G) \perp (a,b) \neq \emptyset$, and that $\delta(out(G) \perp (a,b)) = \{out(G)\}$ otherwise. The is immediate from the first case of the definition of δ. For $\emptyset \subseteq \delta(out(G) \perp (a,b))$ we have already shown that this holds whenever $out(G) \perp (g,h) \neq \emptyset$. The remaining case where $out(G) \perp (g,h) = \emptyset$ is again immediate from the first case of the definition of δ. It only remains to show therefore, that $\delta(out(G) \perp (a,b)) \subseteq out(G) \perp (a,b)$, which is immediate from the second case of the definition of δ.

For $\mathbf{out(G) - (a,b)} = \bigcap \boldsymbol{\delta}(\mathbf{out(G)\perp(a,b)})$. Finally, we need to show that $out(G) - (a,b) = \bigcap \delta(out(G) \perp (a,b))$. There are two cases to consider:

(a) Suppose $out(G) \perp (a,b) = \emptyset$: Then $\bigcap \delta(out(G) \perp (a,b)) = out(G)$, by the definition of δ, whence $out(G) - (a,b) \subseteq out(G) = \bigcap \delta(out(G) \perp (a,b))$ by D-Inclusion. For the converse inclusion, note that $out(G) \perp (a,b) = \emptyset$ implies $\vdash b$. Hence $out(G) - (a, b) = out(G)$ by D-Failure, and therefore $\bigcap \delta(out(G) \perp (a,b)) = out(G) \subseteq out(G) - (a,b)$ as desired.

(b) Suppose $out(G) \perp (a,b) \neq \emptyset$: Since $F \in \delta(out(G) \perp (a,b))$ *iff* $F \in out(G) \perp (a,b)$ and $out(G) - (a,b) \subseteq F$, the inclusion $out(G) - (a,b) \subseteq \bigcap \delta(out(G) \perp (a,b))$ follows immediately from the second case of the definition of δ. For the converse inclusion suppose $(c,d) \in out(G) \setminus out(G) - (a,b)$. We need to find an $F \in \delta(out(G) \perp (a,b))$ such that $(c,d) \notin F$. Since $(c,d) \in out(G) \setminus out(G) - (a,b)$, we have $a \vdash c$ by Input Dependence, whence $(c,d) \in out((out(G) - (a,b)) \cup (c,b))$, by Local Recovery. From lemma 4 it follows that $(c, b \to d) \in out(G) - (a, b)$. By *SI* and D-Closure, therefore, we have $(a, b \to d) \in out(G) - (a, b)$, whence $(a, \neg b \to d) \notin out(G) - (a,b)$ by another application of D-Closure together with *AND*.Therefore $(a, b \vee d) \notin out(G) - (a,b)$ by *WO*, whence $out(G) - (a, b)$ can be extended to a set $F \in out(G) \perp (a, b \vee d)$ by lemma 8. Since $F \in out(G) \perp (a, b \vee d)$ we have $(a, b \vee d) \notin F$, whence $(a,d) \notin F$ by *WO* and $(c,d) \notin F$ by *SI*. Moreover, since $(a, b \vee d) \notin F$ we also have $(a,b) \notin F$, again by *WO*, so $F \in out(G) \perp (a,b)$ by lemma 10. Taking stock we have $F \in out(G) \perp (a,b)$ and $(c,d) \notin F$, so the proof is complete.

This completes the proof.

Proof of theorem 12

Proof: It suffices to find a partial meet derogation operation which yields $\dot{+}$ via the Levi identity. Let δ be defined as follows:

$$\delta(out(G) \perp (a,\neg b)) = \begin{cases} \{out(G)\} & \text{if } out(G) \perp (a,\neg b) \text{is empty, otherwise} \\ \{F \in G \perp (a,\neg b) : (out(G) \dotplus (a,b)) \cap out(G) \subseteq F\} \end{cases}$$

We need to check that δ is a selection function and that it is well-defined:

Well-definedness. We need to show that $\delta(out(G) \perp (a,b)) = \delta(out(G) \perp (c,d))$ whenever $out(G) \perp (a,b) = out(G) \perp (c,d)$. All the limiting cases are similar to theorem 8. For the principal case where $out(G) \perp (a,b) \neq \emptyset$ and $(a, b), (c, d) \in out(G)$ we reason as follows: Since $out(G) \perp (a,b) = out(G) \perp (c,d)$, we have $out((a, b)) = out((c, d))$ by lemma 13. Moreover, since $out(G) \perp (a,b) \neq \emptyset$ we also have $\nvdash b$, whence $a \equiv c$ by lemma 1. By A-Extensionality it follows that $out(G) \dotplus (a,b) = out(G) \dotplus (c,d)$, so $\delta(out(G) \perp (a,b)) = \delta(out(G) \perp (c,d))$ as desired.

***δ* is a selection function**. To show that δ is a selection function, it suffices to show that $\delta(out(G) \perp (a,\neg b)) \neq \emptyset$, whenever $out(G) \perp (a,\neg b) \neq \emptyset$, since, as is easy to check, all other cases are similar to theorem 8. In other words, we need to show, on the assumption that $\delta(out(G) \perp (a,\neg b)) \neq \emptyset$, that there is an $F \in out(G) \perp (a,\neg b)$ such that $(out(G) \dotplus (a,b)) \cap out(G) \subseteq F$. Clearly $(out(G) \dotplus (a,b)) \cap out(G) \subseteq out(G)$, so it suffices to show that $(a,\neg b) \notin out(G) \dotplus (a,b)$, because then $(a,\neg b) \notin (out(G) \dotplus (a,b)) \cap out(G)$ so this set can be extended to an $F \in out(G) \perp (a,\neg b)$ by lemma 8. Suppose for reductio ad absurdum that $(a,\neg b) \in out(G) \dotplus (a,b)$. Then since $(a,b) \in out(G) \dotplus (a,b)$ by A-Success it follows that $(a,f) \in out(G) \dotplus (a,b)$ by *AND*. By A-Consistency therefore $b \vdash f$ whence $\vdash \neg b$. But then $out(G) \perp (a,\neg b) = \emptyset$ contrary to assumption.

Finally we verify the identity

$$out(G) \dotplus (a,b) = out(\bigcap \delta(out(G) \perp (a,\neg b)) \cup (a,b))$$

As a mnemonic device put;

$$out(G) \dotplus_{\delta} (a,b) := out(\bigcap \delta(out(G) \perp (a,\neg b)) \cup (a,b))$$

We thus need to prove that $out(G) \dotplus_{\delta} (a,b) = out(G) \dotplus (a,b)$. We split the proof into two mutually exclusive and jointly exhaustive cases:

(a) Suppose b is inconsistent: Then we have $\vdash \neg b$, so $\bigcap \delta(out(G) \perp (a,\neg b)) = out(G)$ whence $out(G) \dotplus_{\delta} (a,b) = out(G \cup (a,b))$ by the definition of $\dotplus_{\delta}$. It suffices to show, therefore, that $out(G \cup (a,b)) = out(G) \dotplus (a,b)$. The right-in-left is is just A-Inclusion. For the converse we need only show that $out(G) \subseteq out(G) \dotplus (a,b)$, since we then have $out(G) \cup (a,b) \subseteq out(G) \dotplus (a,b)$ by A-Success, whence $out(out(G) \cup (a,b)) \subseteq out(out(G) \dotplus (a,b))$, by monotony for *out*, so $out(G \cup (a,b)) \subseteq out(G) \dotplus (a,b)$ by the closure properties of the *out* operator together with A-Closure. Now, by general set-theory $out(G) \subseteq out(G) \dotplus (a,b)$ whenever $out(G) \setminus out(G) \dotplus (a,b) = \emptyset$, so it suffices to show the latter. Suppose for reduction that $(c,d) \in out(G) \setminus out(G) \dotplus (a,b)$. Then, by A-Relevance there is an $F \subseteq out(G)$ such that $(a,\neg b) \notin out(F)$. However b is inconsistent so $\vdash \neg b$ contradicting $(a,\neg b) \notin out(F)$. Therefore $out(G) \setminus out(G) \dotplus (a,b) = \emptyset$, which completes the case.

(b) In the principal case where b is consistent we argue for each direction of the desired identity separately:

[⇒]: We wish to prove that $out(G)\dot{+}(a,b) \subseteq out(G)\dot{+}_{\delta}(a,b)$. By the definition of δ we have $out(G) \cap (out(G)\dot{+}(a,b)) \subseteq \bigcap \delta(out(G) \perp (a,\neg b))$, from which it follows that $(out(G) \cap (out(G)\dot{+}(a,b))) \cup (a,b) \subseteq \bigcap \delta (out(G) \perp (a,\neg b)) \cup (a,b)$. By monotony for *out*, therefore, we have $out((out(G) \cap (out(G)\dot{+}(a,b))) \cup (a,b)) \subseteq out(\bigcap \delta(out(G) \perp (a,\neg b)) \cup (a,b))$, which, by the definition of $out(G)\dot{+}_{\delta}\ (a,b)$ implies $out((out(G) \cap (out(G)\dot{+}\ (a,b))) \cup (a,b)) \subseteq out(G)\dot{+}_{\delta}(a,b)$. By lemma 18 it thus follows that $out(G)\dot{+}(a,b) \subseteq out(G)\dot{+}_{\delta}(a,b)$, as desired.

[⇐]: We wish to show that $out(G)\dot{+}_{\delta}(a,b) \subseteq out(G)\dot{+}(a,b)$, that is, we want to show that $out(\bigcap \delta(out(G) \perp (a,\neg b)) \cup (a,b)) \subseteq out(G)\dot{+}(a,b)$. By monotony for *out* and A-Closure it suffices to demonstrate that $\bigcap \delta(out(G) \perp (a,\neg b)) \cup (a,b) \subseteq out(G)\dot{+}(a,b)$. Now, $(a,b) \in out(G)\dot{+}(a,b)$, by A-Success, so we need only show that $\bigcap \delta(out(G) \perp (a,\neg b)) \subseteq out(G)\dot{+}(a,b)$. The argument proceeds by contraposition: Suppose $(c,d) \notin out(G)\dot{+}(a,b)$. In the limiting case where $(c,d) \notin out(G)$ we have $(c,d) \notin \bigcap \delta(out(G) \perp (a,\neg b)) \subseteq out(G)$ so we are done. So, suppose $(c, d) \in out(G)$. Then $(c,d) \in out(G) \setminus out(G)\dot{+}(a,b)$. By A-Relevance there is an F such that

1. $out(G) \cap (out(G)\dot{+}(a,b)) \subseteq F \subseteq out(G)$
2. $(a,\neg b) \notin out(F)$, and
3. $(a,\neg b) \in out(F \cup (c,d))$.

By condition 1 and 2 and lemma 8 it follows that F can be extended to an $F' \in out(G) \perp (a,\neg b)$. By condition 3 it follows that $(c,d) \notin F'$, and by 1 again it follows that $out(G) \cap (out(G)\dot{+}(a,b)) \subseteq F'$. Hence, $F' \in \delta(G \perp (a,\neg b))$, whence $(c,d) \notin \bigcap \delta(out(G) \perp (a,\neg b)) = out(G)\dot{+}_{\delta}(a,b)$ as desired.

This completes the proof.

References

Alchourrón CE, Bulygin E (1971) Normative systems. Volume 5 of Library of exact philosophy. Springer

Alchourron CE, Bulygin E (1981) The expressive conception of norms. In: Hilpinen R (ed) New studies in deontic logic. D. Reidel Publishing Company, Dordrecht

Alchourron C, Makinson D (1981) Hierarchies of regulations and their logic. In: Hilpinen R (ed) New studies in deontic logic. D Reidel Publishing Company, Dordrecht

Alchourron CE, Makinson D (1982) On the logic of theory change: contraction functions and their associated revision functions. Theoria 48:14–37

Alchourron CE, Gärdenfors P, Makinson D (1985) On the logic of theory change: partial meet contraction and revision functions. J Symbolic Log 50(2):510–530

Bochman A (2001) A logical theory of nonmonotonic inference and belief change. Springer, Berlin

Boella G, van der Torre L, Verhagen H (2006) Introduction to normative multiagent systems. Comput Math Organ Theory 12:71–79

Boella G, Pigozzi G, van der Torre L (2009) Normative framework for normative system change. In: Proceedings of the 8th international conference on autonomous agents and multiagent systems 1:169–176

Carmo J, Jones AJI (2002) Deontic logic and contrary-to-duties. In: Gabbay D, Guenthner F (eds) Handbook of philosophical logic, vol 8. Kluwer, Netherlands

Davey BA, Priestley HA (2002) Introduction to lattices and order. Cambridge University Press, Cambridge

Dubislav W (1937) Zur Unbegründbarkeit der Forederungsätze. Theoria 3:330–342

Fuhrmann A (1991) Theory contraction through base contraction. J Philos Log 20(2):175–203

Fuhrmann A, Hansson SO (1994) A survey of multiple contractions. J Log Lang Inform 3(1):39–75

Fuhrmann A (1997) An essay on contraction. CSLI Publications, CA

Governatori G, Rotolo A, Sartor G (2005) Temporalised normative positions in defeasible logic. In: Proceedings of the 10th international conference on artificial intelligence and law, pp. 25–34

Governatori G, Rotolo A, Riveret R, Palmirani M, Sartor G (2007) Variants of temporal defeasible logic for modelling norm modifications. In: Proceedings of the 11th international conference on artificial intelligence and law, pp. 155–159

Governatori G, Rotolo A (2008) Changing legal systems: legal abrogations and annulments in defeasible logic. Log J IGPL 18(1):157–194

Gärdenfors P (1988) Knowledge in flux. MIT Press, Cambridge

Gärdenfors P, Rott H (1995) Belief revision. In: Gabbay DM, Hogger CJ, Robinson JA (eds) Handbook of logic in artifcial intelligence and logic programming, volume IV. Oxford University Press, Oxford, pp 35–132

Hansen J (2004) Conflicting imperatives and dyadic deontic logic. In: 7th international workshop on deontic logic in computer science, Springer

Hansson SO (1991) Belief revision without recovery. Studia Log 50:251–260

Hansson SO (1993) Reversing the Levi identity. J Philos Log 22(6):637–669

Hansson SO (1999) A Textbook of Belief Dynamics: Theory Change and Database Updating. Kluwer Academic Publishers, Netherlands

Harper WL (1976a) Ramsey test conditionals and iterated belief change. In: Harper WL, Hooker CA (eds) Foundations of probability theory, statistical inference, and statistical theories of sciences. D. Reidel, Dordrecht, pp 117–135

Harper WL (1976b) Rational conceptual change. PSA 2:462–494

Hilpinen R (2001) Deontic logc. In: Lou G (eds) The Blackwell guide to philosophical logic. Blackwell, Oxford

Jørgensen J (1937) Imperatives and logic. Erkenntnis 7:288–296

Kelsen H (1967) Pure theory of law. Translation from the second (revised and enlarged edition) by Max Knight. University of California press, CA

Kelsen H (1991) General theory of norms. Clarendon press, op. post, Oxford

Shafer-Landau R (2003) Moral realism: a defence. Clarendon press, Oxford

Levi I (1977) Subjunctives, dipositions and chances. Synthese 34:423–455

Levi I (1991) The fixation of belief and its undoing. Cambridge University Press, Cambridge

Makinson D (1987) On the status of the postulate of recovery in the logic of theory change. J Philos Log 16:383–394

Makinson D (1999) On a fundamental problem of deontic logic. In: McNamara P, Prakken H (eds) Norms, logics and information systems. New studies in deontic logic and computer science. IOS Press, Amsterdam, pp 29–53

Makinson DC (2003) Bridges from classical to nonmonotonic Logic. Texts in computing vol 5. King's college press 2005

Makinson D, Gärdenfors P (1988) Revision of knowledge systems and epistemic entrenchment. In: Vardi M (ed) Proceedings of the second conference on theoretical aspects of reasoning about knowledge, Morgan Kaufmann, lk, pp 83–95

Makinson DC, Gärdenfors P (1991) Relations between the logic of theory change and nonmonotonic logic. In: Fuhrmann, Morreu (eds) The logic of theory change, Springer, pp 185–205

Makinson D, van der Torre L (2000) Input/output logics. J Philos Log 29:383–408

Makinson D, van der Torre L (2001) Constraints for input/output logics. J Philos Log 30:155–185

Makinson D, van der Torre L (2003) What is input/output logic. Foundations of the formal sciences II: Applications of mathematical logic in philosophy and linguistics 17:163–174

Makinson D, van der Torre L (2003) Permission from an input/output perspective. J Philos Log 32:391–416

Nebel B (1992) Syntax-based approaches to belief -revision. In: Gärdenfors P (eds) Belief revisison. Cambridge University Press, Cambridge, pp 52–88

Prakken H (1997) Logical tools for modelling legal argument: a study of defeasible reasoning in law. Kluwer, Dodrecht

Prakken H, Sartor G (1997) Argument-based logic programming with defeasible priorities. J Appl Non-classical Log 7:25–75

Ross A (1968) Directives and norms. Routledge and Kegan Paul, London

Rott H (1992) Preferential belief change using generalized epistemic entrenchment. J Log Lang Inform 1:45–78

Segerberg K (2009) Blueprint for a dynamic deontic logic. J Appl Log 7(4)

Stolpe A (2008a) Norms and norm-system dynamics. Ph.D. thesis, Department of Philosophy, University of Bergen, Norway

Stolpe A (2008b) Normative consequence: the problem of keeping it whilst giving it up. In: Proceedings of DEON08, Springer

Stolpe A (2010) A theory of permission based on the notion of derogation. J Appl Log 8:97–113

Stolpe A (2010) Relevance, derogation and permission. In: Proceedings of DEON 2010. Springer

Vranas PBM (2008) New foundations for imperative logic II. Nôus 42(4):529–572

Vranes E (2006) The definition of 'norm conflict' in international law and legal theory. J Int Law 17(2)

Weinberger O (1985) The expressive conception of norms—an impasse for the logic of norms. Law Philos 4(2):165–198

Weitzner DJ, Abelson H, Berners-Lee T, Feigenbaum J, Hendler J, Sussman GJ (2008) Information accountability. Commun ACM 51(6)

Williams M-A, Rott H (eds) (2001) Frontiers in belief revision, vol 22 of Applied logic series. Kluwer Academic Publishers

von Wright GH (1963) Norm and action. Routledge & Kegan Paul, London

von Wright GH (1998) Is and ought. In: Normativity and norms. Clarendon press, Oxford

von Wright GH (1999) Deontic logic—as I see it. In: McNamara P, Prakken H (eds) Norms, logics and information systems. IOS, Amsterdam